GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:16 ; Search time 1701.11 Seconds
(without alignments)
2142.114 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaauucaucg.....uagcgacagcaagcuucugc 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34.6	34.6	642	2	BE147147	BE147147 PM2-HT022
c 2	31.4	31.4	749	9	AG557770	AG557770 Mus muscu
3	31.2	31.2	442	8	AZ955882	AZ955882 2M0222I10
4	30.8	30.8	696	8	B21536	B21536 F21F23-T7 I
c 5	30.6	30.6	461	6	CB736914	CB736914 AMGNNUC:M
6	30.6	30.6	1074	2	BE371824	BE371824 601217621
c 7	30.4	30.4	917	9	AG069018	AG069018 Pan trogl
c 8	30	30.0	725	7	CK016754	CK016754 AGENCOURT
c 9	29.8	29.8	528	8	AQ108017	AQ108017 CIT-HSP-2
c 10	29.6	29.6	578	7	CN553623	CN553623 tae26g08.
c 11	29.6	29.6	617	4	BG838990	BG838990 Gm01_01e0
12	29.6	29.6	626	7	CN623811	CN623811 tae59d12.
13	29.6	29.6	717	7	CN553289	CN553289 tae26g08.
14	29.6	29.6	744	7	CN226782	CN226782 RJB004H01
15	29.6	29.6	745	7	CO421969	CO421969 GGEZHT100
16	29.6	29.6	852	5	BU209960	BU209960 603949596
17	29.6	29.6	877	5	BU399017	BU399017 603534483
18	29.2	29.2	254	4	BM401540	BM401540 JH3C07R S
c 19	29.2	29.2	343	4	BG630802	BG630802 cC-esflcL
20	29.2	29.2	762	7	CF446768	CF446768 EST683113
21	29	29.0	715	8	BZ944697	BZ944697 CH240_121
c 22	28.8	28.8	339	7	CN464219	CN464219 7866.1 Af
c 23	28.8	28.8	532	6	CA249294	CA249294 SCSBFL110
c 24	28.8	28.8	605	9	CE114143	CE114143 tigr-gss-
25	28.8	28.8	739	8	CC115139	CC115139 NDL.81L16
26	28.6	28.6	472	8	AQ955733	AQ955733 LERAF67TR
27	28.6	28.6	513	5	BP051024	BP051024 BP051024
c 28	28.6	28.6	559	6	CB094897	CB094897 hz76e08.b
c 29	28.6	28.6	561	8	BZ891642	BZ891642 CH240_287
c 30	28.4	28.4	580	8	BZ201271	BZ201271 CH230-351
31	28.4	28.4	607	6	CA892404	CA892404 B0172B04-
32	28.4	28.4	629	6	CA890800	CA890800 B0161F04-
33	28.4	28.4	629	6	CA892977	CA892977 B0175H04-
c 34	28.4	28.4	748	9	AG525698	AG525698 Mus muscu
35	28.4	28.4	756	4	BG393101	BG393101 602411369
36	28.4	28.4	818	5	BU708057	BU708057 UI-M-FR0-
c 37	28.4	28.4	832	9	CC491386	CC491386 CH240_325
c 38	28.4	28.4	835	6	CB686108	CB686108 Bn01b_03j
39	28.2	28.2	496	5	BQ531573	BQ531573 APEX2_5_F
40	28.2	28.2	521	5	BQ531501	BQ531501 APEX3_35_
41	28.2	28.2	521	7	CF791691	CF791691 880034 MA
c 42	28.2	28.2	559	4	BG649176	BG649176 EM1_77_E0
c 43	28.2	28.2	598	2	BE363373	BE363373 WS1_62_B0
c 44	28.2	28.2	609	2	BE596636	BE596636 PI1_58_G0
c 45	28.2	28.2	625	6	CA267762	CA267762 SCJLLB207

ALIGNMENTS

RESULT 1
BE147147
LOCUS BE147147 642 bp mRNA linear EST 21-JUN-2000
DEFINITION PM2-HT0224-291099-002-c09 HT0224 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE147147
VERSION BE147147.1 GI:8609871
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 642)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM2-HT0224-291099-002-c09&t3=1999-10-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 415.
FEATURES Location/Qualifiers
source 1. .642
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0224"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 34.6%; Score 34.6; DB 2; Length 642;
Best Local Similarity 37.0%; Pred. No. 0.68;
Matches 30; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

```
Qy          12 AAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
               ||:  |: ||||| : : :|   : | : ::  ::   | ||  :|:||| :||:::::|
Db         150 AATAAATTGAGGCCTGTCTCAGATACTTTTTGGTTTGCAACAAATATAAAATGGTTTTAC 209

Qy          72 CAUAUGCCUUAGCGACAGCAA 92
                 |: :|| ::  | ||| |||
Db         210 AATTTGCGTTTACAACAACAA 230
```

RESULT 2
AG557770/c
LOCUS AG557770 749 bp DNA linear GSS 05-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-476B14.T7, genomic survey sequence.
ACCESSION AG557770
VERSION AG557770.1 GI:48318468
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 749)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institude, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES Location/Qualifiers
source 1. .749
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-476B14.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 31.4%; Score 31.4; DB 9; Length 749;
Best Local Similarity 51.0%; Pred. No. 8.6;
Matches 25; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

```
Qy          41 UUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGACAG 89
               ::||:: |||||||::|:||| :  :  :||||:|:||  :|| | | |
Db         539 TTCCTTAAAACCAGTTATAAAATCTTGCTAGCATATGCAATAGTGTCTG 491
```

RESULT 3
AZ955882
LOCUS AZ955882 442 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0222I10F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0222I10 F, genomic survey sequence.
ACCESSION AZ955882
VERSION AZ955882.1 GI:13827109
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 442)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0222 row: I column: 10
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 442.
FEATURES Location/Qualifiers
source 1. .442
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0222I10"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 31.2%; Score 31.2; DB 8; Length 442;
Best Local Similarity 32.6%; Pred. No. 9.4;
Matches 30; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

```
Qy          7 GGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGU 66
              | |||  ::||:|   |||: :|    |:  |:|::| :: | |  |  :|  ||::|
Db        272 GCAGGCTTTCATCCCAGCATCTGATTCCTGAGTCTTCATTGATAAAATCTAACAATTGAG 331

Qy         67 UUUAGCAUAUGCCUUAGCGACAGCAAGCUUCU 98
              :::    : :   ::  ||| | || | ::|:
Db        332 TTTTTGTTTTTGTTTTTCGAGACCAGGTTTCT 363


RESULT 4
B21536
LOCUS       B21536                  696 bp    DNA     linear   GSS 16-SEP-1997
DEFINITION  F21F23-T7 IGF Arabidopsis thaliana genomic clone F21F23, genomic
            survey sequence.
ACCESSION   B21536
VERSION     B21536.1  GI:2396590
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 696)
  AUTHORS   Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
            Ecker,J.
  TITLE     BAC End Sequences at ATGC
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: F21F23-Sp6
            Contact: Ecker J.
            Arabidopsis Thaliana Genome Center
            University of Pennsylvania
            Dept. of Biology, University of Pennsylvania, Philadelphia, PA
            19104
            Tel: 215-898-9384
            Fax: 215-898-8780
            Email: jecker@atgenome.bio.upenn.edu
            Seq primer: T7
            Class: BAC ends
            High quality sequence start: 129
            High quality sequence stop: 143.
FEATURES             Location/Qualifiers
     source          1. .696
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="F21F23"
                     /sex="hermaphrodite"
                     /clone_lib="IGF"
                     /note="Vector: BeloBACII; Site_1: EcoRI; Site_2: EcoRI;
                     Produced by Thomas Altmann"
ORIGIN

  Query Match             30.8%;  Score 30.8;  DB 8;  Length 696;
  Best Local Similarity   43.9%;  Pred. No. 14;
  Matches   29;  Conservative   15;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          9 AGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUU 68
              || || :||:|||| || |  : ||     : ::||:: | ||||  :|:| |:  |::
Db        306 AGAAAATCATCGAGTCAAAGTTAGAAAGGTTGTTCCTTGACACCACATATAGATGTGTTC 365

Qy         69 UAGCAU 74
              :| ||:
Db        366 TACCAT 371


RESULT 5
CB736914/c
LOCUS       CB736914                461 bp    mRNA    linear   EST 11-APR-2003
DEFINITION  AMGNNUC:MRBE4-00015-C9-A rat brain E15 (10375) Rattus norvegicus
            cDNA clone mrbe4-00015-c9 5', mRNA sequence.
ACCESSION   CB736914
VERSION     CB736914.1  GI:29804127
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00015  row: c  column: 9.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="mrbe4-00015-c9"
                     /tissue_type="brain E15"
                     /clone_lib="rat brain E15 (10375)"
                     /note="Vector: pBCB; Site_1: BstXI; Site_2: Not1; rat
                     brain E15"
ORIGIN

  Query Match             30.6%;  Score 30.6;  DB 6;  Length 461;
  Best Local Similarity   36.7%;  Pred. No. 15;
  Matches   29;  Conservative   19;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              ||||| ::||:|     |: :| |||    :  :  ::|    |||::|:   :  |::
Db        180 GAGGACTTCATCAGTAGATGTGCCGATAAANTTATAGTTCTNGACAGTTATGGGTGAGTT 121

Qy         68 UUAGCAUAUGCCUUAGCGA 86
              : ||||  :|| ::||| |
Db        120 TGAGCAGTTGCATTAGCAA 102


RESULT 6
BE371824
LOCUS       BE371824                1074 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  601217621F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3586661 5',
            mRNA sequence.
ACCESSION   BE371824
VERSION     BE371824.1  GI:9317291
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1074)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8748  row: j  column: 06
            High quality sequence stop: 380.
FEATURES             Location/Qualifiers
     source          1. .1074
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
```

```
                     /clone="IMAGE:3586661"
                     /tissue_type="spontaneous tumor, metastatic to mammary.
                     Stem cell origin."
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu29"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             30.6%;  Score 30.6;  DB 2;  Length 1074;
  Best Local Similarity   44.2%;  Pred. No. 17;
  Matches   34;  Conservative   14;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         19 CGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGC 78
              | |||  :  |  ||| | |:|: ||:  |  ||| ::|: |  :| : : ||||:|: |
Db        280 CCAGGATTCAGCAGACCCTGTCTCCCTGAACCCCACTTATTAGATGCTGTGAGCATATAC 339

Qy         79 CUUAGCGACAGCAAGCU 95
              | : | |||   | ||:
Db        340 CATGGGGACTTGATGCT 356


RESULT 7
AG069018/c
LOCUS       AG069018                917 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-059H17.R, genomic survey sequence.
ACCESSION   AG069018
VERSION     AG069018.1  GI:16620820
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 917)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-059H17.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN

  Query Match             30.4%;  Score 30.4;  DB 9;  Length 917;
  Best Local Similarity   39.1%;  Pred. No. 20;
  Matches   25;  Conservative   18;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         22 GGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUU 81
              || |: :|:|| | || :| :||: |       |::|:|||: | ::::|   :|: ||::
Db        366 GGTATCTGTCGTCCCCCTCCTCCTACGCTGGGGTTATAAATGGATTTTAAATTATACCTT 307

Qy         82 AGCG 85
              ||
Db        306 CCCG 303


RESULT 8
CK016754/c
LOCUS       CK016754                725 bp    mRNA    linear   EST 26-NOV-2003
DEFINITION  AGENCOURT_16544280 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7043617
            5', mRNA sequence.
ACCESSION   CK016754
VERSION     CK016754.1  GI:38541867
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Len Zon, Harvard
             cDNA Library Preparation: Open Biosystems
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14803  row: i  column: 23
            High quality sequence stop: 594.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7043617"
                     /tissue_type="whole body"
                     /lab_host="DH10B"
                     /clone_lib="NIH_ZGC_10"
                     /note="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
                     Bulk tissue was collected from a whole adult individual
                     from the Tuebingen strain. 1st strand cDNA was primed with
                     a Not I - oligo(dT) primer, double-stranded cDNA was
                     cloned into the Not I and EcoRV sites of pExpress-1.
                     Library was size-selected for >1 kb fragments. A
                     normalized version of this library is also available
                     (NIH_ZGC_7). Library was constructed by Open Biosystems
                     (Huntsville, AL)."
ORIGIN

  Query Match             30.0%;  Score 30;  DB 7;  Length 725;
  Best Local Similarity   44.3%;  Pred. No. 26;
  Matches   31;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          2 GGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAA 61
              ||||:| | || ::|| | || ||:     |  :|||:  :| :: ||| ||  :|:| |
Db        663 GGAGTGTAAGATTTCAGCAAGACATTGTGGGCATCCGTAGTCTTTGAAAGCATATATATA 604

Qy         62 UUGGUUUUAG 71
              :  |:  :||
Db        603 TAAGTAGTAG 594
```

```
RESULT 9
AQ108017/c
LOCUS       AQ108017                528 bp    DNA     linear   GSS 29-AUG-1998
DEFINITION  CIT-HSP-2374K6.TF CIT-HSP Homo sapiens genomic clone 2374K6,
            genomic survey sequence.
ACCESSION   AQ108017
VERSION     AQ108017.1  GI:3484196
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CIT-HSP-2374K6.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2374K6"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN

  Query Match             29.8%;  Score 29.8;  DB 8;  Length 528;
  Best Local Similarity   36.9%;  Pred. No. 29;
  Matches   24;  Conservative   19;  Mismatches   22;  Indels    0;  Gaps    0;

Qy         13 AUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGC 72
              |:: |:  ||| |:|: : |||: ||: ::| ::  | | |  :| |||::|  :::
Db        411 ATTAATACAGGTATATATTGACTGCGTTTTCTTTGGACCAAAATACAAATTGAATTTTAT 352

Qy         73 AUAUG 77
              | |:|
Db        351 ACATG 347


RESULT 10
CN553623/c
LOCUS       CN553623                578 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  tae26g08.y1 Hydra EST Darmstadt I Hydra magnipapillata cDNA 5'
            similar to TR:Q9VJZ2 Q9VJZ2 CG6565 PROTEIN. ;, mRNA sequence.
ACCESSION   CN553623
VERSION     CN553623.1  GI:46962927
KEYWORDS    EST.
SOURCE      Hydra magnipapillata
  ORGANISM  Hydra magnipapillata
            Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
            Hydridae; Hydra.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
            Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
            Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,
            Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
            Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
  TITLE     WashU Hydra EST Project
  JOURNAL   Unpublished (2002)
COMMENT     Other_ESTs: tae26g08.x1
            Contact: Hans Bode
            WashU Hydra EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by Corina Guder / GATC Konstanz, Germany
            Library materials provided by Thomas Holstein / Molecular Cell
            Biology, TUD, Darmstadt DNA sequencing by: Washington University
            Genome Sequencing Center For information on obtaining a clone
            please contact: Hans Bode (hrbode@uci.edu)
            Seq primer: -40UP
            High quality sequence stop: 565.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Hydra magnipapillata"
                     /mol_type="mRNA"
                     /strain="sf-1 mutant of strain 105"
                     /db_xref="taxon:6085"
                     /lab_host="TransforMax EC100 (Epicentre), T1 Phage
                     resistant electrocompetent cells"
                     /clone_lib="Hydra EST Darmstadt I"
                     /note="Vector: pBluescript II SK (+); Site_1: NotI;
                     Site_2: EcoRI"
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 7;  Length 578;
  Best Local Similarity   40.8%;  Pred. No. 34;
  Matches   31;  Conservative   16;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         21 AGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCU 80
              | | |:| |   ||:  |:   ||    ||| ||:: : ||:: |::: | ||: :|  :
Db        453 AAGGATACGGGCACTTTGTGACCCAATGAACAAGTTGTGAATTAGTTTCATCATTTGAGT 394

Qy         81 UAGCGACAGCAAGCUU 96
              | ||| | |||| ::
Db        393 CAACGAGATCAAGTTT 378


RESULT 11
BG838990/c
LOCUS       BG838990                617 bp    mRNA    linear   EST 25-MAY-2001
DEFINITION  Gm01_01e04_F Gm01_AAFC_ECORC_Glycine_max_cold_stressed_leaves
            Glycine max cDNA clone Gm01_01e04, mRNA sequence.
ACCESSION   BG838990
VERSION     BG838990.1  GI:14205312
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Singh,J.A., Robert,L.S., Lu,B., Zhu,L., De Moors,A., Couroux,P.,
            Harris,L.J., Hattori,J.I., Ouellet,T., Sprott,D. and Tinker,N.A.
  TITLE     Expressed Sequence Tags from Cold-Stressed Glycine max (soybean)
            Leaves
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
```

```
            Email: singhja@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1. .617
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /cultivar="Maple Arrow"
                     /db_xref="taxon:3847"
                     /clone="Gm01_01e04"
                     /tissue_type="Leaves"
                     /dev_stage="15-day seedlings"
                     /clone_lib="Gm01_AAFC_ECORC_Glycine_max_cold_stressed_leav
                     es"
                     /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: EcoRI;
                     Site_2: XhoI; Plants were grown 12 days from seeds,
                     treated at 2oC for 3 days, then  leaves were collected.
                     Library was prepared using Unizap ZAP-cDNA  synthesis kit
                     / Packaged Gigapack III Gold."
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 4;  Length 617;
  Best Local Similarity   42.4%;  Pred. No. 35;
  Matches   36;  Conservative   14;  Mismatches   35;  Indels    0;  Gaps    0;

Qy         14 UUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCA 73
              : ||    | ||  |: : ||| |  :| :|  : ||| |||::| || :  |::  | ||
Db        123 TGCAGGCAAGCCAATTTTGACCCACTCCTCAATAAAATCAGTTAAAATTANGTTCCATCA 64

Qy         74 UAUGCCUUAGCGACAGCAAGCUUCU 98
              :|:|     |    |||| || |  |:
Db         63 TATGGGGAAAAAACAGAAACCCCCT 39


RESULT 12
CN623811
LOCUS       CN623811                 626 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  tae59d12.x1 Hydra EST Darmstadt I Hydra magnipapillata cDNA 3'
            similar to TR:Q9VJZ2 Q9VJZ2 CG6565 PROTEIN. ;, mRNA sequence.
ACCESSION   CN623811
VERSION     CN623811.1  GI:47134888
KEYWORDS    EST.
SOURCE      Hydra magnipapillata
  ORGANISM  Hydra magnipapillata
            Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
            Hydridae; Hydra.
REFERENCE   1  (bases 1 to 626)
  AUTHORS   Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
            Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
            Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,
            Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
            Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
  TITLE     WashU Hydra EST Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hans Bode
            WashU Hydra EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by Corina Guder / GATC Konstanz, Germany
            Library materials provided by Thomas Holstein / Molecular Cell
            Biology, TUD, Darmstadt DNA sequencing by: Washington University
            Genome Sequencing Center For information on obtaining a clone
            please contact: Hans Bode (hrbode@uci.edu)
            Seq primer: degenerate primer
            High quality sequence stop: 626.
FEATURES             Location/Qualifiers
     source          1. .626
                     /organism="Hydra magnipapillata"
                     /mol_type="mRNA"
                     /strain="sf-1 mutant of strain 105"
                     /db_xref="taxon:6085"
                     /lab_host="TransforMax EC100 (Epicentre), T1 Phage
                     resistant electrocompetent cells"
                     /clone_lib="Hydra EST Darmstadt I"
                     /note="Vector: pBluescript II SK (+); Site_1: NotI;
                     Site_2: EcoRI"
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 7;  Length 626;
  Best Local Similarity   40.8%;  Pred. No. 35;
  Matches   31;  Conservative   16;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         21 AGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCU 80
              ||| |:|:|   ||:  |:   ||    ||| ||:: : || : |::: | ||: :|  :
Db        377 AGGGATATGGGCACTTTGTGACCCAATGAACAAGTTGTGAACTAGTTTCATCATTTGAGT 436

Qy         81 UAGCGACAGCAAGCUU 96
              | | | | |||| ::
Db        437 CAACAAGATCAAGTTT 452


RESULT 13
CN553289
LOCUS       CN553289                 717 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  tae26g08.x1 Hydra EST Darmstadt I Hydra magnipapillata cDNA 3'
            similar to TR:Q9VJZ2 Q9VJZ2 CG6565 PROTEIN. ;, mRNA sequence.
ACCESSION   CN553289
VERSION     CN553289.1  GI:46962593
KEYWORDS    EST.
SOURCE      Hydra magnipapillata
  ORGANISM  Hydra magnipapillata
            Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
            Hydridae; Hydra.
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,
            Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,
            Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,
            Gibbons,M., Ritter,E., Bennett,J., Ronko,I., Tsagareishvili,R.,
            Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.
  TITLE     WashU Hydra EST Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hans Bode
            WashU Hydra EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by Corina Guder / GATC Konstanz, Germany
            Library materials provided by Thomas Holstein / Molecular Cell
            Biology, TUD, Darmstadt DNA sequencing by: Washington University
            Genome Sequencing Center For information on obtaining a clone
            please contact: Hans Bode (hrbode@uci.edu)
            Seq primer: degenerate primer.
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="Hydra magnipapillata"
                     /mol_type="mRNA"
                     /strain="sf-1 mutant of strain 105"
                     /db_xref="taxon:6085"
                     /lab_host="TransforMax EC100 (Epicentre), T1 Phage
                     resistant electrocompetent cells"
                     /clone_lib="Hydra EST Darmstadt I"
                     /note="Vector: pBluescript II SK (+); Site_1: NotI;
                     Site_2: EcoRI"
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 7;  Length 717;
  Best Local Similarity   40.8%;  Pred. No. 36;
  Matches   31;  Conservative   16;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         21 AGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCU 80
              | | |:| |   ||:  |:   ||    ||| ||:: : ||:: |::: | ||: :|  :
Db        466 AAGGATACGGGCACTTTGTGACCCAATGAACAAGTTGTGAATTAGTTTCATCATTTGAGT 525
```

```
Qy          81 UAGCGACAGCAAGCUU 96
               | ||| | |||| ::
Db         526 CAACGAGATCAAGTTT 541


RESULT 14
CN226782
LOCUS       CN226782                744 bp    mRNA    linear   EST 09-APR-2004
DEFINITION  RJB004H01.ab1 RJtestis Gallus gallus cDNA 5', mRNA sequence.
ACCESSION   CN226782
VERSION     CN226782.1  GI:46330526
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and
            Lundeberg,J.
  TITLE     EST analysis of brain and testis cDNA libraries from White leghorn
            and Red Jungle Fowl
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Peter Savolainen
            Department of Biotechnology
            Royal Institute of Technology, KTH
            SE-106 91 Stockholm, SWEDEN
            Tel: +46 (0)8 5537 8481
            Fax: +46 (0)8 5537 8335
            Email: Peter.Savolainen@biotech.kth.se
            Seq primer: M13 reverse primer.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Red junglefowl"
                     /db_xref="taxon:9031"
                     /sex="male"
                     /lab_host="ElectroMAX DH10B (Invitrogen)"
                     /clone_lib="RJtestis"
                     /note="Organ: testis; Vector: pSPORT-1; Site_1: Hind III;
                     Site_2: EcoRI; The cDNA libraries were created with the
                     Superscript Plasmid System  (Invitrogen)."
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 7;  Length 744;
  Best Local Similarity   55.8%;  Pred. No. 36;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACC 52
              |||||:  | |  ::||:||||||| |:|    |:|||:|: || :|  |||
Db        491 GGGAGTTTAAGGCTTCATCGAGGCAAATGGGAGCTCCGTCTCCCATCGCACC 542


RESULT 15
CO421969
LOCUS       CO421969                745 bp    mRNA    linear   EST 02-JUL-2004
DEFINITION  GGEZHT1005A10.g HT1 Gallus gallus cDNA clone GGEZHT1005A10, mRNA
            sequence.
ACCESSION   CO421969
VERSION     CO421969.1  GI:49638217
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   Silva,C.S., Jorge,E.C., Patricio,M., Ledur,M.C. and Coutinho,L.L.
  TITLE     Discovery of new genes expressed in the chicken pituitary and
            hypothalamus
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Clarissa S. Silva
            Laboratory of Animal Biotecnology, Dep. of Animal Production
            ESALQ - University of Sao Paulo
            Av. Padua Dias, 11, Piracicaba, SP, 13418-900, Brazil
            Tel: 55 19 3429 4434
            Fax: 55 19 3429 4285
            Email: cssilva@esalq.usp.br and llcoutin@esalq.usp.br
            PCR PRimers
            BACKWARD: T7.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
                     /clone="GGEZHT1005A10"
                     /tissue_type="pituitary and hypothalamus"
                     /dev_stage="21 days old"
                     /lab_host="DH10B"
                     /clone_lib="HT1"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
                     cDNA library was constructed with the SuperScript Plasmid
                     System with Gateway Tecnology kit (Invitrogen), following
                     manufacture's protocols. Plasmid DNA was purified using a
                     modified alkaline lysis method. Sequencing reactions were
                     conducted using the kit Big Dye Terminator Cycle
                     Sequencing Ready Reaction (Applied Biosystems) according
                     to the manufacturer's  recommendations. Clones were
                     sequenced by the 5' end with T7 primer. Sequencing
                     reactions were analyzed on ABI Prism 3100 Genetic Analyzer
                     (Applied Biosystems). The quality and clustering of the
                     ESTs were analyzed using the softwares Phred/Cap3. Only
                     EST sequences with Phred quality greater than 20 and at
                     least 150 bp were considered for clustering."
ORIGIN

  Query Match             29.6%;  Score 29.6;  DB 7;  Length 745;
  Best Local Similarity   55.8%;  Pred. No. 36;
  Matches   29;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACC 52
              |||||:  | |  ::||:||||||| |:|    |:|||:|: || :|  |||
Db        254 GGGAGTTTAAGGCTTCATCGAGGCAAATGGGAGCTCCGTCTCCCATCGCACC 305


Search completed: November 30, 2004, 12:52:13
Job time : 1709.11 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:15 ; Search time 1093.89 Seconds
(without alignments)
4323.081 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaauucaucg.....uagcgacagcaagcuucugc 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		32.2	32.2	199415	10	AC084053	AC084053 Mus Muscu
c 2		31.2	31.2	201197	10	AC125279	AC125279 Mus muscu
3		30.8	30.8	187606	2	AC110847	AC110847 Rattus no
c 4		30.8	30.8	221981	2	AC103165	AC103165 Rattus no
c 5		30.8	30.8	231814	2	AC114165	AC114165 Rattus no
6		30.8	30.8	233823	2	AC110697	AC110697 Rattus no
c 7		30.8	30.8	238172	2	AC099144	AC099144 Rattus no
c 8		30.8	30.8	250740	2	AC129051	AC129051 Rattus no
c 9		30.6	30.6	748	6	CQ451733	CQ451733 Sequence
10		30.6	30.6	15776	6	CQ363795	CQ363795 Sequence
11		30.6	30.6	110000	1	AE017283_15	Continuation (16 o
c 12		30.4	30.4	217	4	AF522912	AF522912 Tachyglos
c 13		30.4	30.4	217	4	AF522914	AF522914 Tachyglos
c 14		30.4	30.4	217	4	AF522916	AF522916 Tachyglos
c 15		30.4	30.4	366	6	AR318058	AR318058 Sequence
c 16		30.4	30.4	3185	4	AY194920	AY194920 Tachyglos
c 17		30.4	30.4	46077	2	AC137289	AC137289 Rattus no
18		30.4	30.4	58143	2	AC137177	AC137177 Rattus no
c 19		30.4	30.4	207260	2	AC134220	AC134220 Rattus no
20		30.4	30.4	213033	2	AC131482	AC131482 Rattus no
c 21		30.4	30.4	218157	2	AC139588	AC139588 Rattus no
22		30.4	30.4	272400	2	AC110146	AC110146 Rattus no
23		30.4	30.4	283299	2	AC112576	AC112576 Rattus no
c 24		30.2	30.2	191004	2	CR536601	CR536601 Ornithorh
25		30.2	30.2	229522	2	AC109911	AC109911 Rattus no
c 26		30.2	30.2	235937	2	AC131223	AC131223 Rattus no
27		30.2	30.2	240755	2	AC095368	AC095368 Rattus no
28		30	30.0	213257	2	AC133102	AC133102 Mus muscu
29		29.8	29.8	4285	6	AR372457	AR372457 Sequence
30		29.8	29.8	4285	8	AF057708	AF057708 Populus b
c 31		29.8	29.8	184147	9	AC103805	AC103805 Homo sapi
c 32		29.8	29.8	191844	2	AC148836	AC148836 Pan trogl
c 33		29.8	29.8	202287	2	AC145887	AC145887 Pan trogl
34		29.8	29.8	215994	9	AC067941	AC067941 Homo sapi
c 35		29.8	29.8	223143	2	CR450730	CR450730 Danio rer
36		29.6	29.6	23393	2	AC018216	AC018216 Drosophil
37		29.6	29.6	50481	2	AC117541	AC117541 Mus muscu
38		29.6	29.6	79826	3	AC002443	AC002443 Drosophil
c 39		29.6	29.6	110000	2	AC110832_3	Continuation (4 of
40		29.6	29.6	110000	2	AC112373_04	Continuation (5 of
41		29.6	29.6	156806	3	AC093047	AC093047 Drosophil
42		29.6	29.6	182816	5	AC145913	AC145913 Gallus ga
43		29.6	29.6	192132	3	AC093440	AC093440 Drosophil
c 44		29.6	29.6	234627	2	AC106118	AC106118 Rattus no
45		29.6	29.6	237005	2	AC137343	AC137343 Rattus no

ALIGNMENTS

RESULT 1
AC084053
LOCUS AC084053 199415 bp DNA linear ROD 05-SEP-2002
DEFINITION Mus Musculus Strain C57BL6/J chromosome 3 BAC, RP23-158F1, complete sequence.
ACCESSION AC084053
VERSION AC084053.6 GI:22725947
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 199415)
AUTHORS Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE High Throughput Mouse Sequencing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 199415)
AUTHORS Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-2000) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA
REFERENCE 3 (bases 1 to 199415)
AUTHORS Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
REFERENCE 4 (bases 1 to 199415)
AUTHORS Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M., Goltz,J., Haider,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J., Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA

COMMENT On Sep 5, 2002 this sequence version replaced gi:22002115.
----------Genome Center:
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: http://www.hpcgg.org/Sequence/mouse.html
Contact: hpgc@mendel.mgh.harvard.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550).
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double strandedsequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation asLow Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

----------Summary Statistics
Center project name: ADW
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 199415
Fraction of Phrap value < 40: 0.000326
Error Rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

---------- Distribution of Quality < 40 Bases:

----------.

FEATURES Location/Qualifiers
source 1. .199415
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="3"
/clone="RP23-158F1"
repeat_region complement(557. .1289)
/rpt_family="L1MA7"
repeat_region 1304. .1331
/rpt_family="AT_rich"
repeat_region complement(1392. .1889)
/rpt_family="Lx"
repeat_region complement(1888. .3694)
/rpt_family="Lx"
repeat_region 3696. .4251
/rpt_family="L1F"
repeat_region 4252. .4441
/rpt_family="B2_Mm1"
repeat_region 4442. .4577
/rpt_family="L1F"
repeat_region complement(5123. .5507)
/rpt_family="Lx"
repeat_region complement(5512. .6982)
/rpt_family="L1VL2"
repeat_region complement(7008. .7281)
/rpt_family="L1"
repeat_region complement(7527. .8154)
/rpt_family="L1_MM"
repeat_region 9064. .9276
/rpt_family="Lx5"
repeat_region 9257. .9816
/rpt_family="L1VL2"
repeat_region 9796. .9882
/rpt_family="Lx5"
repeat_region 9883. .9934
/rpt_family="(CA)n"
repeat_region 9935. .10658
/rpt_family="Lx5"
repeat_region 10679. .10955
/rpt_family="Lx5"
repeat_region 10957. .10985
/rpt_family="(CAAA)n"
repeat_region complement(11152. .11712)
/rpt_family="Lx"
repeat_region 11713. .12270
/rpt_family="Lx5"
repeat_region 12269. .13913
/rpt_family="L1_MM"
repeat_region complement(13902. .14155)
/rpt_family="L1"
repeat_region 14155. .14258
/rpt_family="L1_MM"
repeat_region 14256. .14355
/rpt_family="Lx5"
repeat_region 14364. .14675
/rpt_family="Lx5"
repeat_region 14738. .14809
/rpt_family="CT-rich"
repeat_region complement(15461. .15814)
/rpt_family="MLT2B3"
repeat_region complement(15831. .15915)
/rpt_family="MTE"
repeat_region 16451. .16511
/rpt_family="ID5"
repeat_region 17106. .17159
/rpt_family="(TCTA)n"
repeat_region 17356. .17382
/rpt_family="(TCC)n"
repeat_region 18211. .18244
/rpt_family="(CAAA)n"
repeat_region 18325. .18841
/rpt_family="MLT2B3"
repeat_region 18892. .18915
/rpt_family="AT_rich"
repeat_region 19461. .20516
/rpt_family="Lx4"
repeat_region complement(20510. .22297)
/rpt_family="L1"
repeat_region 22298. .23888
/rpt_family="L1_MM"

```
     repeat_region   23882. .24118
                     /rpt_family="Lx4"
     repeat_region   complement(24228. .24641)
                     /rpt_family="Lx6"
     repeat_region   complement(24821. .25092)
                     /rpt_family="Lx3"
     repeat_region   25183. .25208
                     /rpt_family="(TAAA)n"
     repeat_region   25210. .25277
                     /rpt_family="(GAAA)n"
     repeat_region   complement(25288. .25661)
                     /rpt_family="MTC"
     repeat_region   complement(26114. .32348)
                     /rpt_family="L1F"

  Query Match             32.2%;  Score 32.2;  DB 10;  Length 199415;
  Best Local Similarity   49.1%;  Pred. No. 8.5;
  Matches   26;  Conservative   14;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         37 CGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGACAG 89
              | : ::||:: |||||||::|:||| :  :  :|||:|:||  :|| | | |
Db      63221 CTTATTCCTTAAAACCAGTTATAAAATCTTGCTAGCATATGCAATAGTGTCTG 63273


RESULT 2
AC125279/c
LOCUS       AC125279               201197 bp    DNA     linear   ROD 05-NOV-2003
DEFINITION  Mus musculus BAC clone RP23-433F5 from 8, complete sequence.
ACCESSION   AC125279
VERSION     AC125279.1  GI:21536173
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 201197)
  AUTHORS   Trani,L. and Cotton,M.
  TITLE     The sequence of Mus musculus BAC clone RP23-433F5
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 201197)
  AUTHORS   Wilson,R.
  TITLE     Sequencing of Mus musculus
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 201197)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 201197)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     ---------- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
            ---------- Summary Statistics
              Center project name: M_BA0433F05
            ----------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES                Location/Qualifiers
     source             1. .201197
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        /chromosome="8"
                        /map="8"
                        /clone="RP23-433F5"
                        /clone_lib="RPCI-23"
     repeat_region      171. .557
                        /rpt_family="MaLR"
     repeat_region      558. .1352
                        /rpt_family="MaLR"
     repeat_region      1353. .1499
                        /rpt_family="Alu"
     repeat_region      1500. .1886
                        /rpt_family="MaLR"
     repeat_region      1887. .2101
                        /rpt_family="MaLR"
     repeat_region      2103. .2485
                        /rpt_family="L1"
     repeat_region      2792. .2918
                        /rpt_family="Alu"
     repeat_region      2916. .3397
                        /rpt_family="RMER6B"
     repeat_region      5516. .5712
                        /rpt_family="MaLR"
     repeat_region      6071. .6162
                        /rpt_family="ERVK"
     repeat_region      6163. .6556
                        /rpt_family="MaLR"
     repeat_region      6557. .6980
                        /rpt_family="ERVK"
     repeat_region      7137. .7220
                        /rpt_family="MIR"
     repeat_region      7970. .8262
                        /rpt_family="L1"
     repeat_region      8734. .9125
                        /rpt_family="L1"
     repeat_region      9194. .10473
                        /rpt_family="L1"
     repeat_region      10485. .10619
                        /rpt_family="MER1_type"
     repeat_region      10573. .10940
                        /rpt_family="ERVK"
     repeat_region      10673. .11051
                        /rpt_family="RMER17C"
     repeat_region      11200. .12140
                        /rpt_family="L1"
     repeat_region      12674. .12714
                        /rpt_family="L1"
     repeat_region      12715. .12907
                        /rpt_family="B2"
     repeat_region      12908. .13742
                        /rpt_family="L1"
```

```
repeat_region   14760. .15436
                /rpt_family="L1"
repeat_region   16028. .16131
                /rpt_family="MaLR"
repeat_region   17836. .18263
                /rpt_family="L1"
repeat_region   18352. .18653
                /rpt_family="ERVK"
repeat_region   18655. .19078
                /rpt_family="ERVK"
repeat_region   19375. .19942
                /rpt_family="L1"
repeat_region   21030. .21117
                /rpt_family="MaLR"
repeat_region   22308. .22444
                /rpt_family="MIR"
repeat_region   22775. .22988
                /rpt_family="MER1_type"
repeat_region   22991. .23088
                /rpt_family="L1"
repeat_region   23274. .23681
                /rpt_family="MaLR"
repeat_region   24784. .24971
                /rpt_family="B2"
repeat_region   25314. .25486
                /rpt_family="MaLR"
repeat_region   26145. .26223
                /rpt_family="RAL_RN"
repeat_region   26473. .26619
                /rpt_family="Alu"
repeat_region   26817. .27001
                /rpt_family="B2"
tRNA            complement(26924. .26996)
                /product="tRNA-Ser"
                /note="Likely pseudogene (HMM Sc=39.59 / Sec struct
                Sc=-11.21)"
repeat_region   27016. .27079
                /rpt_family="Alu"
repeat_region   28465. .28604
                /rpt_family="Alu"
repeat_region   28818. .29133
                /rpt_family="MaLR"
repeat_region   29251. .29389
                /rpt_family="MIR"
repeat_region   29546. .29687
                /rpt_family="MaLR"
repeat_region   30593. .30737
                /rpt_family="L2"
repeat_region   32073. .32437
                /rpt_family="MaLR"
repeat_region   33137. .33276
                /rpt_family="Alu"
repeat_region   33402. .33547
                /rpt_family="Alu"
repeat_region   35207. .35290
                /rpt_family="MIR"
repeat_region   35653. .35830
                /rpt_family="MaLR"
repeat_region   35831. .36039
                /rpt_family="B2"
repeat_region   36040. .36232
                /rpt_family="MaLR"
repeat_region   36324. .36549
                /rpt_family="B4"
repeat_region   36709. .36788
                /rpt_family="B4"
repeat_region   36795. .36891
                /rpt_family="ERVL"
repeat_region   36862. .37099
                /rpt_family="MaLR"
repeat_region   39693. .39806
                /rpt_family="MER1_type"
repeat_region   40903. .41102
                /rpt_family="MER2_type"
repeat_region   41131. .41296
                /rpt_family="MER2_type"
repeat_region   41519. .41626
                /rpt_family="L2"
repeat_region   42585. .42730
                /rpt_family="Alu"

  Query Match            31.2%;  Score 31.2;  DB 10;  Length 201197;
  Best Local Similarity  32.6%;  Pred. No. 19;
  Matches   30;  Conservative   24;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          7 GGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGU 66
              | |||  ::||:|   |||: :|    |:  |:|::| :: | |  |  :|  ||::|
Db       2984 GCAGGCTTTCATCCCAGCATCTGATTCCTGAGTCTTCATTGATAAAATCTAACAATTGAG 2925

Qy         67 UUUAGCAUAUGCCUUAGCGACAGCAAGCUUCU 98
              :::    : :   ::  ||| | || | ::|:
Db       2924 TTTTTGTTTTTGTTTTCGAGACCAGGTTTCT 2893


RESULT 3
AC110847
LOCUS       AC110847              187606 bp    DNA     linear   HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-42M10, *** SEQUENCING IN PROGRESS
            ***, 13 unordered pieces.
ACCESSION   AC110847
VERSION     AC110847.6  GI:25006792
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 187606)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwackelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
```

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 187606)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (16-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 187606)

AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Nov 15, 2002 this sequence version replaced gi:23321762.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
-------------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
Center project name: GRXS
Center clone name: CH230-42M10
-------------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 118671 bases at least Q40
Consensus quality: 124131 bases at least Q30
Consensus quality: 127839 bases at least Q20
Estimated insert size: 122165; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31167: contig of 31167 bp in length
* 31168 31267: gap of unknown length
* 31268 77383: contig of 46116 bp in length
* 77384 77483: gap of unknown length
* 77484 168911: contig of 91428 bp in length
* 168912 169011: gap of unknown length
* 169012 170421: contig of 1410 bp in length
* 170422 170521: gap of unknown length
* 170522 172029: contig of 1508 bp in length
* 172030 172129: gap of unknown length
* 172130 173707: contig of 1578 bp in length
* 173708 173807: gap of unknown length
* 173808 175104: contig of 1297 bp in length
* 175105 175204: gap of unknown length
* 175205 176766: contig of 1562 bp in length
* 176767 176866: gap of unknown length
* 176867 178988: contig of 2122 bp in length
* 178989 179088: gap of unknown length
* 179089 180430: contig of 1342 bp in length
* 180431 180530: gap of unknown length
* 180531 182073: contig of 1543 bp in length
* 182074 182173: gap of unknown length
* 182174 184491: contig of 2318 bp in length
* 184492 184591: gap of unknown length
* 184592 187606: contig of 3015 bp in length.

FEATURES Location/Qualifiers
source 1. .187606
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-42M10"
misc_feature 13985. .15402
/note="wgs_contig"
misc_feature 28318. .29829
/note="wgs_contig"
misc_feature 77484. .79244
/note="wgs_contig"
misc_feature 95087. .96968
/note="wgs_contig"
misc_feature 131603. .132754
/note="wgs_contig"
misc_feature 145131. .150091
/note="wgs_contig"
misc_feature 167411. .168911
/note="wgs_contig"

ORIGIN

Query Match 30.8%; Score 30.8; DB 2; Length 187606;
Best Local Similarity 34.7%; Pred. No. 25;
Matches 34; Conservative 22; Mismatches 42; Indels 0; Gaps 0;

```
Qy          2 GGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAA 61
              | || | || |  :||:   |   :|: :   | | |: ::  :  |||    ::|: |
Db     164680 GTAGAGCAGAATATCATTAGGAGTTATTTAATCACTGTTTTTTTCTAAAGTCATTATTAT 164739

Qy         62 UUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUG 99
              :: |::::  | :| | | : | | || |||| ::| |
Db     164740 TTAGTTTGCCCTAGGTCCTTGGGCCATCAAGTTTCAG 164777
```

RESULT 4
AC103165/c
LOCUS AC103165 221981 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-194N19, WORKING DRAFT SEQUENCE.
ACCESSION AC103165
VERSION AC103165.6 GI:30578614
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 221981)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

```
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 221981)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 221981)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:25087265.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GJJH
                Center clone name: CH230-194N19
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 210155 bases at least Q40
                Consensus quality: 214132 bases at least Q30
                Consensus quality: 216741 bases at least Q20
                Estimated insert size: 222472; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   221981: contig of 221981 bp in length.
FEATURES            Location/Qualifiers
     source         1. .221981
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-194N19"
     misc_feature   1. .1454
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   39959. .40695
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH339559"
ORIGIN

  Query Match            30.8%;  Score 30.8;  DB 2;  Length 221981;
  Best Local Similarity  37.8%;  Pred. No. 26;
  Matches   28;  Conservative   19;  Mismatches   27;  Indels    0;  Gaps    0;

Qy         26 UAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCG 85
              : :|  || :| |: ::| ::|  || |  : :| |: ||:: :|||   |  ::   |
Db      43994 TTTGGGGAATCTGTTTTCTTTCCTACTACCTGTATATAGGTTCTAGCAATGGAATTCAGG 43935

Qy         86 ACAGCAAGCUUCUG 99
                || || ||::|:|
Db      43934 TCATCAGGCTTCTG 43921


RESULT 5
AC114165/c
LOCUS       AC114165              231814 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-165B14, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC114165
VERSION     AC114165.4  GI:30579575
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 231814)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
```

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 231814)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 231814)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 13, 2003 this sequence version replaced gi:23664775.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GLOV
Center clone name: CH230-165B14
---------- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 219004 bases at least Q40
Consensus quality: 222024 bases at least Q30
Consensus quality: 224099 bases at least Q20
Estimated insert size: 229718; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 231814: contig of 231814 bp in length.

FEATURES Location/Qualifiers
source 1. .231814
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-165B14"
misc_feature 1. .1811
/note="wgs_contig"
misc_feature 1
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH272064"
misc_feature 93963. .96094
/note="wgs_contig"
misc_feature 119152. .121641
/note="wgs_contig"
misc_feature complement(227278. .227563)
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH272063"
misc_feature 228398. .231814
/note="wgs_end_extension
clone_end:T7"

ORIGIN

Query Match 30.8%; Score 30.8; DB 2; Length 231814;
Best Local Similarity 34.7%; Pred. No. 26;
Matches 34; Conservative 22; Mismatches 42; Indels 0; Gaps 0;

```
Qy          2 GGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAA 61
              | || | || |  :||:  |  :|: :  | | |: ::  :  |||   ::|: |
Db     230424 GTAGAGCAGAATATCATTAGGAGTTATTTAATCACTGTTTTTTTCTAAAGTCATTATTAT 230365

Qy         62 UUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUG 99
              :: |::::  | :| | | : | | || |||| ::| |
Db     230364 TTAGTTTGCCCTAAGTCCTTGGGCCATCAAGTTTCAG 230327
```

RESULT 6
AC110697
LOCUS AC110697 233823 bp DNA linear HTG 23-NOV-2002
DEFINITION Rattus norvegicus clone CH230-43H7, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC110697
VERSION AC110697.5 GI:25188485
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 233823)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 233823)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 233823)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 23, 2002 this sequence version replaced gi:23267238.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GRXR
Center clone name: CH230-43H7
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 212738 bases at least Q40
Consensus quality: 215401 bases at least Q30
Consensus quality: 217067 bases at least Q20
Estimated insert size: 219523; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 233823: contig of 233823 bp in length.

FEATURES Location/Qualifiers
source 1. .233823
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-43H7"
misc_feature 1. .1276
/note="wgs_end_extension
clone_end:T7"
misc_feature 1327. .2821
/note="wgs_end_extension
clone_end:T7"
misc_feature 7344. .8215
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH264888"
misc_feature 85851. .88514
/note="wgs_contig"
misc_feature complement(232161. .232332)
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH264890"

ORIGIN

Query Match 30.8%; Score 30.8; DB 2; Length 233823;
Best Local Similarity 34.7%; Pred. No. 26;
Matches 34; Conservative 22; Mismatches 42; Indels 0; Gaps 0;

Qy 2 GGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAA 61
Db 230212 GTAGAGCAGAATATCATTAGGAGTTATTTAATCACTGTTTTTTTCTAAAGTCATTATTAT 230271

Qy 62 UUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUG 99
Db 230272 TTAGTTTTGCCCTAGGTCCTTGGGCCATCAAGTTTCAG 230309

```
RESULT 7
AC099144/c
LOCUS       AC099144               238172 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-68F7, WORKING DRAFT SEQUENCE.
ACCESSION   AC099144
VERSION     AC099144.5  GI:30522119
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 238172)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 238172)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 238172)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:23264509.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GGYP
                Center clone name: CH230-68F7
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 220106 bases at least Q40
                Consensus quality: 222963 bases at least Q30
                Consensus quality: 224727 bases at least Q20
                Estimated insert size: 230633; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *         1   238172: contig of 238172 bp in length.
FEATURES             Location/Qualifiers
     source          1. .238172
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-68F7"
     misc_feature    1. .1721
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    8345. .8845
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH351992"
     misc_feature    234541. .234906
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH351956"
     misc_feature    236172. .238172
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

  Query Match             30.8%;  Score 30.8;  DB 2;  Length 238172;
  Best Local Similarity   43.1%;  Pred. No. 26;
  Matches   25;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         23 GCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCU 80
```

```
                 |||:|: :||| :|| :|:: | : |||||| ::  ||  :||: ::|  | |:  |:
Db     129774 GCATATTTCGATTCCTTCTTGCCTTAAACCAATTTGAACGTGGTCTTACAAAATCTCT 129717


RESULT 8
AC129051/c
LOCUS       AC129051               250740 bp    DNA     linear   HTG 20-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-89P7, *** SEQUENCING IN PROGRESS ***,
            3 unordered pieces.
ACCESSION   AC129051
VERSION     AC129051.3  GI:25139158
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 250740)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 250740)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 250740)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:23683080.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GQOW
                Center clone name: CH230-89P7
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 217247 bases at least Q40
                Consensus quality: 222917 bases at least Q30
                Consensus quality: 226268 bases at least Q20
                Estimated insert size: 225713; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   247784: contig of 247784 bp in length
            *  247785   247884: gap of unknown length
            *  247885   249178: contig of 1294 bp in length
            *  249179   249278: gap of unknown length
            *  249279   250740: contig of 1462 bp in length.
FEATURES             Location/Qualifiers
     source          1. .250740
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-89P7"
     misc_feature    1. .631
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BH343662"
     misc_feature    37160. .38616
                     /note="wgs_contig"
     misc_feature    39370. .40812
                     /note="wgs_contig"
     misc_feature    152186. .153864
                     /note="wgs_contig"
     misc_feature    162672. .164912
                     /note="wgs_contig"
     misc_feature    164963. .166515
                     /note="wgs_contig"
```

```
     misc_feature    179234. .181122
                     /note="wgs_contig"
ORIGIN

  Query Match             30.8%;  Score 30.8;  DB 2;  Length 250740;
  Best Local Similarity   43.1%;  Pred. No. 26;
  Matches   25;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         23 GCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCU 80
              |||:|: :||| :|| :|:: | : |||||| ::  ||  :||: ::|  | |:  |:
Db       4082 GCATATTTCGATTCCTTCTTGCCTTAAACCAATTTGAACGTGGTCTTACAAAATCTCT 4025


RESULT 9
CQ451733/c
LOCUS       CQ451733                748 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 17493 from Patent WO0192523.
ACCESSION   CQ451733
VERSION     CQ451733.1  GI:41420097
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.D.
  TITLE     Human polynucleotides and polypeptides encoded thereby
  JOURNAL   Patent: WO 0192523-A 17493 06-DEC-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .748
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             30.6%;  Score 30.6;  DB 6;  Length 748;
  Best Local Similarity   46.8%;  Pred. No. 20;
  Matches   36;  Conservative   12;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              : | ||   |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db        317 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 258

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db        257 CCCTGGGCACGGGAAGC 241


RESULT 10
CQ363795
LOCUS       CQ363795              15776 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 78 from Patent WO0181581.
ACCESSION   CQ363795
VERSION     CQ363795.1  GI:41300489
KEYWORDS    .
SOURCE      Propionibacterium acnes
  ORGANISM  Propionibacterium acnes
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Propionibacterineae; Propionibacteriaceae; Propionibacterium.
REFERENCE   1
  AUTHORS   Skeiky,Y.A., Persing,D.H., Mitcham,J.L., Wang,S.S., Bhatia,A.,
            L'Maisonneuve,J.F., Zhang,Y., Jen,S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of acne
            vulgaris
  JOURNAL   Patent: WO 0181581-A 78 01-NOV-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .15776
                     /organism="Propionibacterium acnes"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1747"
ORIGIN

  Query Match             30.6%;  Score 30.6;  DB 6;  Length 15776;
  Best Local Similarity   46.8%;  Pred. No. 25;
  Matches   36;  Conservative   12;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              : | ||   |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db       8427 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 8486

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db       8487 CCCTGGGCACGGGAAGC 8503


RESULT 11
AE017283_15
WPCOMMENT
  Sequence split into 26 fragments  LOCUS AE017283 Accession AE017283
     Fragment Name      Begin     End
     AE017283_00            1    110000
     AE017283_01       100001    210000
     AE017283_02       200001    310000
     AE017283_03       300001    410000
     AE017283_04       400001    510000
     AE017283_05       500001    610000
     AE017283_06       600001    710000
     AE017283_07       700001    810000
     AE017283_08       800001    910000
     AE017283_09       900001   1010000
     AE017283_10      1000001   1110000
     AE017283_11      1100001   1210000
     AE017283_12      1200001   1310000
     AE017283_13      1300001   1410000
     AE017283_14      1400001   1510000
     AE017283_15      1500001   1610000
     AE017283_16      1600001   1710000
     AE017283_17      1700001   1810000
     AE017283_18      1800001   1910000
     AE017283_19      1900001   2010000
     AE017283_20      2000001   2110000
     AE017283_21      2100001   2210000
     AE017283_22      2200001   2310000
     AE017283_23      2300001   2410000
     AE017283_24      2400001   2510000
     AE017283_25      2500001   2560265
Continuation (16 of 26) of AE017283 from base 1500001 (AE017283 Propionibacterium acnes P

  Query Match             30.6%;  Score 30.6;  DB 1;  Length 110000;
  Best Local Similarity   46.8%;  Pred. No. 29;
  Matches   36;  Conservative   12;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              : | ||   |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db      34666 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 34725

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db      34726 CCCTGGGCACGGGAAGC 34742


RESULT 12
AF522912/c
LOCUS       AF522912                217 bp    DNA     linear   MAM 22-MAY-2003
DEFINITION  Tachyglossus aculeatus clone TaG7 type I interferon gene, partial
            cds.
ACCESSION   AF522912
VERSION     AF522912.1  GI:27451581
KEYWORDS    .
SOURCE      Tachyglossus aculeatus (Australian echidna)
  ORGANISM  Tachyglossus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Monotremata; Tachyglossidae; Tachyglossus.
REFERENCE   1  (bases 1 to 217)
```

```
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     A survey of type I interferons from a marsupial and monotreme:
            implications for the evolution of the type I interferon gene family
            in mammals
  JOURNAL   Cytokine 21 (3), 105-119 (2003)
  MEDLINE   22583464
   PUBMED   12697149
REFERENCE   2  (bases 1 to 217)
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2002) School of Science, Food and Horticulture,
            University of Western Sydney, Locked Bag 1797, Penrith South DC,
            NSW 1797, Australia
FEATURES             Location/Qualifiers
     source          1. .217
                     /organism="Tachyglossus aculeatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9261"
                     /clone="TaG7"
     mRNA            <1. .>217
                     /product="type I interferon"
     CDS             <1. .>217
                     /codon_start=2
                     /product="type I interferon"
                     /protein_id="AAO14986.1"
                     /db_xref="GI:27451582"
                     /translation="NIFSRNHSQTGWDEAIVEKFLHGVHQEMVWLGLFLEEEMGWENS
                     TLRRDISLHIKSYFKRMMDYLKGRNYSS"
ORIGIN

  Query Match             30.4%;  Score 30.4;  DB 4;  Length 217;
  Best Local Similarity   41.7%;  Pred. No. 21;
  Matches   30;  Conservative   16;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          6 UGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGG 65
              :|  ||| : ||: |||| |: : :||||:  | |::| : | | ||||::    | :|
Db         78 TGATGGACTCCATGGAGGAATTTCTCGACTATGGCTTCATCCCAGCCAGTTTGGGAGTGA 19

Qy         66 UUUUAGCAUAUG 77
              :::  ||  | |
Db         18 TTTCTGCTGAAG 7


RESULT 13
AF522914/c
LOCUS       AF522914                 217 bp    DNA     linear   MAM 22-MAY-2003
DEFINITION  Tachyglossus aculeatus clone TaG11 type I interferon gene, partial
            cds.
ACCESSION   AF522914
VERSION     AF522914.1  GI:27451585
KEYWORDS    .
SOURCE      Tachyglossus aculeatus (Australian echidna)
  ORGANISM  Tachyglossus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Monotremata; Tachyglossidae; Tachyglossus.
REFERENCE   1  (bases 1 to 217)
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     A survey of type I interferons from a marsupial and monotreme:
            implications for the evolution of the type I interferon gene family
            in mammals
  JOURNAL   Cytokine 21 (3), 105-119 (2003)
  MEDLINE   22583464
   PUBMED   12697149
REFERENCE   2  (bases 1 to 217)
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2002) School of Science, Food and Horticulture,
            University of Western Sydney, Locked Bag 1797, Penrith South DC,
            NSW 1797, Australia
FEATURES             Location/Qualifiers
     source          1. .217
                     /organism="Tachyglossus aculeatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9261"
                     /clone="TaG11"
     mRNA            <1. .>217
                     /product="type I interferon"
     CDS             <1. .>217
                     /codon_start=2
                     /product="type I interferon"
                     /protein_id="AAO14988.1"
                     /db_xref="GI:27451586"
                     /translation="NIFSRNHSQTGWDEAIVEKFLHGVHQEMVWLELFLEEEMGWENS
                     TLRRDISLHIKSYFKRMMDYLKGRNYSS"
ORIGIN

  Query Match             30.4%;  Score 30.4;  DB 4;  Length 217;
  Best Local Similarity   41.7%;  Pred. No. 21;
  Matches   30;  Conservative   16;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          6 UGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGG 65
              :|  ||| : ||: |||| |: : :||||:  | |::| : | | ||||::    | :|
Db         78 TGATGGACTCCATGGAGGAATTTCTCGACTATGGCTTCATCCCAGCCAGTTTGGGAGTGA 19

Qy         66 UUUUAGCAUAUG 77
              :::  ||  | |
Db         18 TTTCTGCTGAAG 7


RESULT 14
AF522916/c
LOCUS       AF522916                 217 bp    DNA     linear   MAM 22-MAY-2003
DEFINITION  Tachyglossus aculeatus clone TaG31 type I interferon gene, partial
            cds.
ACCESSION   AF522916
VERSION     AF522916.1  GI:27451589
KEYWORDS    .
SOURCE      Tachyglossus aculeatus (Australian echidna)
  ORGANISM  Tachyglossus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Monotremata; Tachyglossidae; Tachyglossus.
REFERENCE   1  (bases 1 to 217)
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     A survey of type I interferons from a marsupial and monotreme:
            implications for the evolution of the type I interferon gene family
            in mammals
  JOURNAL   Cytokine 21 (3), 105-119 (2003)
  MEDLINE   22583464
   PUBMED   12697149
REFERENCE   2  (bases 1 to 217)
  AUTHORS   Harrison,G.A., Young,L.J., Watson,C.M., Miska,K.B., Miller,R.D. and
            Deane,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2002) School of Science, Food and Horticulture,
            University of Western Sydney, Locked Bag 1797, Penrith South DC,
            NSW 1797, Australia
FEATURES             Location/Qualifiers
     source          1. .217
                     /organism="Tachyglossus aculeatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9261"
                     /clone="TaG31"
     mRNA            <1. .>217
                     /product="type I interferon"
     CDS             <1. .>217
                     /codon_start=2
                     /product="type I interferon"
                     /protein_id="AAO14990.1"
                     /db_xref="GI:27451590"
                     /translation="NIFSRNHSQTGWDEAIVEKFLHGVHQEMVWLELFLEEEMGWENS
                     TLRRDISLHIKSYFKRMMDYLKGRNYSS"
```

```
ORIGIN

  Query Match             30.4%;  Score 30.4;  DB 4;  Length 217;
  Best Local Similarity   41.7%;  Pred. No. 21;
  Matches   30;  Conservative   16;  Mismatches   26;  Indels    0;  Gaps    0;

Qy          6 UGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGG 65
              :|  ||| : ||: |||| |: : :||||:  | |::| : | | ||||::    | :|
Db         78 TGATGGACTCCATGGAGGAATTTCTCGACTATGGCTTCATCCCAGCCAGTTTGGGAGTGA 19

Qy         66 UUUUAGCAUAUG 77
              :::  ||  | |
Db         18 TTTCTGCTGAAG 7


RESULT 15
AR318058/c
LOCUS       AR318058                366 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 608 from patent US 6562958.
ACCESSION   AR318058
VERSION     AR318058.1  GI:33699161
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Breton,G. and Bush,D.
  TITLE     Nucleic acid and amino acid sequences relating to Acinetobacter
            baumannii for diagnostics and therapeutics
  JOURNAL   Patent: US 6562958-A 608 13-MAY-2003;
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             30.4%;  Score 30.4;  DB 6;  Length 366;
  Best Local Similarity   39.8%;  Pred. No. 22;
  Matches   35;  Conservative   17;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||| |:  | |||:  | |     |:|:|:||   | |:  :  : |||||  ||: |:|
Db        101 GGGTGTTTATGAATATAGCTTTCTATATGTCGCACCAGTAATTGTGCAAACAAGTAATAG 42

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACA 88
                :||:::  || :|:    :||   ||
Db         41 TCTGGTTTATGCGTATATTCTAGATTCA 14


Search completed: November 30, 2004, 11:52:20
Job time : 1100.89 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:11 ; Search time 232.222 Seconds
(without alignments)
2260.516 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaauucaucg.....uagcgacagcaagcuucugc 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		100	100.0	100	3	AAZ99048	Aaz99048 RNA aptam
2		100	100.0	100	12	ADJ10063	Adj10063 Synthetic
3		100	100.0	100	12	ADJ10064	Adj10064 Synthetic
4		100	100.0	200	12	ADJ10065	Adj10065 Synthetic
5		90	90.0	90	3	AAZ99050	Aaz99050 RNA aptam
6		84.4	84.4	98	12	ADJ10053	Adj10053 RNA aptam
7		80	80.0	80	3	AAZ99051	Aaz99051 RNA aptam
8		50.6	50.6	147	12	ADJ10066	Adj10066 Synthetic
9		45	45.0	45	3	AAZ99076	Aaz99076 Ras targe
c 10		30.6	30.6	748	3	AAC75113	Aac75113 Human ORF
c 11		30.6	30.6	748	6	ABN24508	Abn24508 Human ORF
12		30.6	30.6	15776	4	AAS59583	Aas59583 Propionib
13		30.6	30.6	15776	8	ACF64512	Acf64512 Propionib
c 14		30.4	30.4	366	9	ADA29321	Ada29321 DNA encod
15		30	30.0	98	3	AAZ99049	Aaz99049 RNA aptam
16		29.8	29.8	4192	3	AAZ57942	Aaz57942 Poplar fl
17		29.8	29.8	4285	4	AAF85391	Aaf85391 Nucleotid
18		29.8	29.8	4285	6	ABK88484	Abk88484 Poplar pr
19		29.8	29.8	4285	9	ACA62517	Aca62517 Poplar ho
20		29	29.0	29	3	AAZ99084	Aaz99084 Ras targe
c 21		29	29.0	29	12	ADJ10055	Adj10055 Synthetic
c 22		28.8	28.8	675	8	ACA52761	Aca52761 Prokaryot
c 23		28.6	28.6	2000	6	ABZ15103	Abz15103 Arabidops
c 24		28.2	28.2	1161	5	ABV24321	Abv24321 Human pro
c 25		28.2	28.2	1665	2	AAX58401	Aax58401 Vicia sat
c 26		28	28.0	29	12	ADJ10058	Adj10058 PCR prime
c 27		28	28.0	29	12	ADJ10057	Adj10057 PCR prime
c 28		28	28.0	29	12	ADJ10056	Adj10056 PCR prime
c 29		28	28.0	110000	6	ABX08336_01	Continuation (2 of
c 30		28	28.0	110000	12	ADJ25985_01	Continuation (2 of
c 31		28	28.0	110000	12	ADN97989_01	Continuation (2 of
c 32		28	28.0	110000	12	ADO50281_01	Continuation (2 of
33		27.6	27.6	732	4	AAI94947	Aai94947 Human neu
34		27.6	27.6	732	8	ABT42823	Abt42823 Human neu
c 35		27.4	27.4	3791	2	AAT63575	Aat63575 Chicken b
c 36		27	27.0	27	3	AAZ99087	Aaz99087 Ras targe
c 37		27	27.0	29	12	ADJ10059	Adj10059 PCR prime
c 38		27	27.0	208	3	AAC17425	Aac17425 Human sec
39		27	27.0	522	3	AAA69206	Aaa69206 Bacteriop
40		27	27.0	56506	3	AAA69168	Aaa69168 Bacteriop
c 41		26.8	26.8	415	3	AAC21986	Aac21986 Human sec
42		26.8	26.8	1116	5	AAH51835	Aah51835 HIV prote
43		26.8	26.8	1116	12	ADH44700	Adh44700 HIV prote
c 44		26.6	26.6	48995	9	ADA02579	Ada02579 Mouse Rel
c 45		26.6	26.6	48995	10	ADB72317	Adb72317 Mouse Rel

ALIGNMENTS

RESULT 1
AAZ99048
ID AAZ99048 standard; RNA; 100 BP.
XX
AC AAZ99048;
XX
DT 21-JUN-2000 (first entry)
XX
DE RNA aptamer #25 for binding Ras target protein.
XX
KW Ras target protein; malignant tumour; signal transmission regulation;
KW cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200009684-A1.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-JP004399.
XX
PR 14-AUG-1998; 98JP-00242596.
PR 24-NOV-1998; 98JP-00333284.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Yokoyama S, Hirao I, Sakamoto K;
XX
DR WPI; 2000-224330/19.
XX
PT Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT target protein like Raf-1, useful in drug compositions to treat and
PT diagnose malignant tumors and inflammation.
XX
PS Claim 6; Page 41; 59pp; Japanese.
XX
CC The invention relates to novel nucleic acids which bind specifically to
CC the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-Z99051)
CC based on these sequences are useful in the treatment and diagnosis of
CC malignant tumours and inflammation. The nucleic acids can be used to
CC formulate medicinal compositions that are useful in the treatment of
CC malignant tumours and inflammation and for disease diagnosis by binding
CC specifically to Ras target protein and regulating transmission of signal
CC causing proliferation or differentiation of cells

```
XX
SQ   Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

  Query Match             100.0%;  Score 100;  DB 3;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-25;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
              ||||||||||||||||||||||||||||||||||||||||
Db         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100


RESULT 2
ADJ10063
ID   ADJ10063 standard; DNA; 100 BP.
XX
AC   ADJ10063;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Synthetic RNA 9A (100-mer).
XX
KW   ss; unnatural base; 5'-substituted-2-oxo(1H)- pyridin-3-yl;
KW   photoreactive group; biotin; fluorescent molecule;
KW   6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS   Synthetic.
XX
PN   WO2004007713-A1.
XX
PD   22-JAN-2004.
XX
PF   28-FEB-2003; 2003WO-JP002342.
XX
PR   17-JUL-2002; 2002JP-00208568.
XX
PA   (RIKE ) RIKEN KK.
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Hirao I,  Yokoyama S,  Hirao M,  Mitsui T;
XX
DR   WPI; 2004-122944/12.
XX
PT   Nucleosides or nucleotides comprising novel base particularly 5-
PT   substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT   functional nucleic acids as reagents or therapeutics in treating
PT   diseases.
XX
PS   Disclosure; Fig 10; 78pp; Japanese.
XX
CC   This invention relates to nucleosides or nucleotides that a novel
CC   unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group
CC   as a base. Specifically, it refers to nucleic acid molecules integrated
CC   with a base substituted at the 5-position with either a photoreactive
CC   group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group,
CC   biotin or a derivative thereof or a fluorescent molecule selected from
CC   fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC   derivatives thereof. The present invention describes preparing nucleic
CC   acids by carrying out transcription, replication or reverse transcription
CC   with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC   base which is applied as template for the integration of any of the
CC   nucleotides into its complementary position. Accordingly, the nucleosides
CC   and nucleotides are useful in developing functional nucleic acids
CC   including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC   therapeutics in treating diseases by gene therapy. The produced nucleic
CC   acids are chemically stable with improved ease of amplification,
CC   replication and transcription for their preparation. This polynucleotide
CC   sequence is a synthetic RNA 9A aptamer (100-mer) given in an
CC   exemplification of the invention.
XX
SQ   Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

  Query Match             100.0%;  Score 100;  DB 12;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-25;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
              ||||||||||||||||||||||||||||||||||||||||
Db         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100


RESULT 3
ADJ10064
ID   ADJ10064 standard; DNA; 100 BP.
XX
AC   ADJ10064;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Synthetic RNA 9A(5Iy87) aptamer (100-mer).
XX
KW   ss; unnatural base; 5'-substituted-2-oxo(1H)- pyridin-3-yl;
KW   photoreactive group; biotin; fluorescent molecule;
KW   6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   87
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "OTHER= 5Iy"
XX
PN   WO2004007713-A1.
XX
PD   22-JAN-2004.
XX
PF   28-FEB-2003; 2003WO-JP002342.
XX
PR   17-JUL-2002; 2002JP-00208568.
XX
PA   (RIKE ) RIKEN KK.
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Hirao I,  Yokoyama S,  Hirao M,  Mitsui T;
XX
DR   WPI; 2004-122944/12.
XX
PT   Nucleosides or nucleotides comprising novel base particularly 5-
PT   substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT   functional nucleic acids as reagents or therapeutics in treating
PT   diseases.
XX
PS   Disclosure; Fig 10; 78pp; Japanese.
XX
CC   This invention relates to nucleosides or nucleotides that a novel
CC   unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group
CC   as a base. Specifically, it refers to nucleic acid molecules integrated
CC   with a base substituted at the 5-position with either a photoreactive
CC   group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group,
CC   biotin or a derivative thereof or a fluorescent molecule selected from
CC   fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC   derivatives thereof. The present invention describes preparing nucleic
CC   acids by carrying out transcription, replication or reverse transcription
CC   with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC   base which is applied as template for the integration of any of the
```

CC nucleotides into its complementary position. Accordingly, the nucleosides
CC and nucleotides are useful in developing functional nucleic acids
CC including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC therapeutics in treating diseases by gene therapy. The produced nucleic
CC acids are chemically stable with improved ease of amplification,
CC replication and transcription for their preparation. This polynucleotide
CC sequence is a synthetic RNA 9A(5Iy87) aptamer (100-mer) given in an
CC exemplification of the invention.
XX
SQ Sequence 100 BP; 25 A; 22 C; 24 G; 0 T; 29 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.4e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
Db 1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy 61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
Db 61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100

RESULT 4
ADJ10065
ID ADJ10065 standard; DNA; 200 BP.
XX
AC ADJ10065;
XX
DT 17-JUN-2004 (first entry)
XX
DE Synthetic RNA 2x9A aptamer (200-mer).
XX
KW ss; unnatural base; 5'-substituted-2-oxo(1H)- pyridin-3-yl;
KW photoreactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS Synthetic.
XX
PN WO2004007713-A1.
XX
PD 22-JAN-2004.
XX
PF 28-FEB-2003; 2003WO-JP002342.
XX
PR 17-JUL-2002; 2002JP-00208568.
XX
PA (RIKE) RIKEN KK.
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Hirao I, Yokoyama S, Hirao M, Mitsui T;
XX
DR WPI; 2004-122944/12.
XX
PT Nucleosides or nucleotides comprising novel base particularly 5-
PT substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT functional nucleic acids as reagents or therapeutics in treating
PT diseases.
XX
PS Disclosure; Fig 10; 78pp; Japanese.
XX
CC This invention relates to nucleosides or nucleotides that a novel
CC unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group
CC as a base. Specifically, it refers to nucleic acid molecules integrated
CC with a base substituted at the 5-position with either a photoreactive
CC group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group,
CC biotin or a derivative thereof or a fluorescent molecule selected from
CC fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC derivatives thereof. The present invention describes preparing nucleic
CC acids by carrying out transcription, replication or reverse transcription
CC with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC base which is applied as template for the integration of any of the
CC nucleotides into its complementary position. Accordingly, the nucleosides
CC and nucleotides are useful in developing functional nucleic acids
CC including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC therapeutics in treating diseases by gene therapy. The produced nucleic
CC acids are chemically stable with improved ease of amplification,
CC replication and transcription for their preparation. This polynucleotide
CC sequence is a synthetic RNA 2x9A aptamer (200-mer) given in an
CC exemplification of the invention.
XX
SQ Sequence 200 BP; 50 A; 44 C; 48 G; 0 T; 58 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 200;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
Db 1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy 61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
Db 61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100

RESULT 5
AAZ99050
ID AAZ99050 standard; RNA; 90 BP.
XX
AC AAZ99050;
XX
DT 21-JUN-2000 (first entry)
XX
DE RNA aptamer #27 for binding Ras target protein.
XX
KW Ras target protein; malignant tumour; signal transmission regulation;
KW cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200009684-A1.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-JP004399.
XX
PR 14-AUG-1998; 98JP-00242596.
PR 24-NOV-1998; 98JP-00333284.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Yokoyama S, Hirao I, Sakamoto K;
XX
DR WPI; 2000-224330/19.
XX
PT Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT target protein like Raf-1, useful in drug compositions to treat and
PT diagnose malignant tumors and inflammation.
XX
PS Claim 6; Page 42; 59pp; Japanese.
XX
CC The invention relates to novel nucleic acids which bind specifically to
CC the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-Z99051)
CC based on these sequences are useful in the treatment and diagnosis of
CC malignant tumours and inflammation. The nucleic acids can be used to
CC formulate medicinal compositions that are useful in the treatment of
CC malignant tumours and inflammation and for disease diagnosis by binding
CC specifically to Ras target protein and regulating transmission of signal
CC causing proliferation or differentiation of cells
XX
SQ Sequence 90 BP; 23 A; 19 C; 22 G; 0 T; 26 U; 0 Other;

Query Match 90.0%; Score 90; DB 3; Length 90;

```
  Best Local Similarity   100.0%;  Pred. No. 8e-22;
  Matches   90;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGC 90
              |||||||||||||||||||||||||||||||
Db         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGC 90


RESULT 6
ADJ10053
ID   ADJ10053 standard; RNA; 98 BP.
XX
AC   ADJ10053;
XX
DT   17-JUN-2004  (first entry)
XX
DE   RNA aptamer (RNA9A) SeqID 1.
XX
KW   ss; unnatural base; 5'-substituted-2-oxo(1H)- pyridin-3-yl;
KW   photoreactive group; biotin; fluorescent molecule;
KW   6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS   Unidentified.
XX
PN   WO2004007713-A1.
XX
PD   22-JAN-2004.
XX
PF   28-FEB-2003; 2003WO-JP002342.
XX
PR   17-JUL-2002; 2002JP-00208568.
XX
PA   (RIKE ) RIKEN KK.
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Hirao I,  Yokoyama S,  Hirao M,  Mitsui T;
XX
DR   WPI; 2004-122944/12.
XX
PT   Nucleosides or nucleotides comprising novel base particularly 5-
PT   substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT   functional nucleic acids as reagents or therapeutics in treating
PT   diseases.
XX
PS   Example 4; SEQ ID NO 1; 78pp; Japanese.
XX
CC   This invention relates to nucleosides or nucleotides that a novel
CC   unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group
CC   as a base. Specifically, it refers to nucleic acid molecules integrated
CC   with a base substituted at the 5-position with either a photoreactive
CC   group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group,
CC   biotin or a derivative thereof or a fluorescent molecule selected from
CC   fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC   derivatives thereof. The present invention describes preparing nucleic
CC   acids by carrying out transcription, replication or reverse transcription
CC   with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC   base which is applied as template for the integration of any of the
CC   nucleotides into its complementary position. Accordingly, the nucleosides
CC   and nucleotides are useful in developing functional nucleic acids
CC   including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC   therapeutics in treating diseases by gene therapy. The produced nucleic
CC   acids are chemically stable with improved ease of amplification,
CC   replication and transcription for their preparation. This polynucleotide
CC   sequence is an RNA aptamer (RNA9A) given in an exemplification of the
CC   invention.
XX
SQ   Sequence 98 BP; 24 A; 21 C; 24 G; 0 T; 29 U; 0 Other;

  Query Match             84.4%;  Score 84.4;  DB 12;  Length 98;
  Best Local Similarity   97.0%;  Pred. No. 7.9e-20;
  Matches   97;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGUGGAGGAAUUC--UGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 58

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
              ||||||||||||||||||||||||||||||||||||||||
Db         59 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 98


RESULT 7
AAZ99051
ID   AAZ99051 standard; RNA; 80 BP.
XX
AC   AAZ99051;
XX
DT   21-JUN-2000  (first entry)
XX
DE   RNA aptamer #28 for binding Ras target protein.
XX
KW   Ras target protein; malignant tumour; signal transmission regulation;
KW   cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009684-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1999;   99WO-JP004399.
XX
PR   14-AUG-1998;   98JP-00242596.
PR   24-NOV-1998;   98JP-00333284.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Yokoyama S,  Hirao I,  Sakamoto K;
XX
DR   WPI; 2000-224330/19.
XX
PT   Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT   target protein like Raf-1, useful in drug compositions to treat and
PT   diagnose malignant tumors and inflammation.
XX
PS   Claim 6; Page 42; 59pp; Japanese.
XX
CC   The invention relates to novel nucleic acids which bind specifically to
CC   the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-Z99051)
CC   based on these sequences are useful in the treatment and diagnosis of
CC   malignant tumours and inflammation. The nucleic acids can be used to
CC   formulate medicinal compositions that are useful in the treatment of
CC   malignant tumours and inflammation and for disease diagnosis by binding
CC   specifically to Ras target protein and regulating transmission of signal
CC   causing proliferation or differentiation of cells
XX
SQ   Sequence 80 BP; 20 A; 16 C; 19 G; 0 T; 25 U; 0 Other;

  Query Match             80.0%;  Score 80;  DB 3;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-18;
  Matches   80;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60

Qy         61 AUUGGUUUUAGCAUAUGCCU 80
              ||||||||||||||||||||
Db         61 AUUGGUUUUAGCAUAUGCCU 80
```

RESULT 8
ADJ10066
ID ADJ10066 standard; DNA; 147 BP.
XX
AC ADJ10066;
XX
DT 17-JUN-2004 (first entry)
XX
DE Synthetic RNA 0C aptamer (100-mer).
XX
KW ss; unnatural base; 5'-substituted-2-oxo(1H)- pyridin-3-yl;
KW photoreactive group; biotin; fluorescent molecule;
KW 6-substituted-2-aminopurin-9-yl; aptamer; gene therapy.
XX
OS Synthetic.
XX
PN WO2004007713-A1.
XX
PD 22-JAN-2004.
XX
PF 28-FEB-2003; 2003WO-JP002342.
XX
PR 17-JUL-2002; 2002JP-00208568.
XX
PA (RIKE) RIKEN KK.
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Hirao I, Yokoyama S, Hirao M, Mitsui T;
XX
DR WPI; 2004-122944/12.
XX
PT Nucleosides or nucleotides comprising novel base particularly 5-
PT substituted-2-oxo(1H)-pyridin-3-yl group as base, useful in developing
PT functional nucleic acids as reagents or therapeutics in treating
PT diseases.
XX
PS Disclosure; Fig 10; 78pp; Japanese.
XX
CC This invention relates to nucleosides or nucleotides that a novel
CC unnatural base, namely the 5'-substituted-2-oxo(1H)- pyridin-3-yl group
CC as a base. Specifically, it refers to nucleic acid molecules integrated
CC with a base substituted at the 5-position with either a photoreactive
CC group (e.g. iodine or bromine), an alkenyl, alkynyl or amino group,
CC biotin or a derivative thereof or a fluorescent molecule selected from
CC fluorescein, 6-carboxyfluorescein, tetramethyl-7-carboxyrhodamine or
CC derivatives thereof. The present invention describes preparing nucleic
CC acids by carrying out transcription, replication or reverse transcription
CC with a nucleic acid containing a 6-substituted-2-aminopurin-9-yl group as
CC base which is applied as template for the integration of any of the
CC nucleotides into its complementary position. Accordingly, the nucleosides
CC and nucleotides are useful in developing functional nucleic acids
CC including antisense DNAs and RNAs, ribozymes and aptamers as reagents or
CC therapeutics in treating diseases by gene therapy. The produced nucleic
CC acids are chemically stable with improved ease of amplification,
CC replication and transcription for their preparation. This polynucleotide
CC sequence is a synthetic RNA 0C aptamer (100-mer) given in an
CC exemplification of the invention.
XX
SQ Sequence 147 BP; 32 A; 34 C; 37 G; 0 T; 44 U; 0 Other;

Query Match 50.6%; Score 50.6; DB 12; Length 147;
Best Local Similarity 75.2%; Pred. No. 8e-08;
Matches 76; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

```
Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGU-UAUA 59
              |||||||||||||||||||||||||| ||   || |  |||| |||         | |||
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUCUGGGAACCCUAUCUUGCUUUUGGUAGCUGUAUU 60

Qy         60 AAUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
                 |   ||   |||||||||||||||||||||||||||||
Db         61 CACCUGUAACAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 101
```

RESULT 9
AAZ99076
ID AAZ99076 standard; RNA; 45 BP.
XX
AC AAZ99076;
XX
DT 21-JUN-2000 (first entry)
XX
DE Ras target protein-RNA binding sequence #1 for generating primers.
XX
KW Ras target protein; malignant tumour; signal transmission regulation;
KW cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200009684-A1.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-JP004399.
XX
PR 14-AUG-1998; 98JP-00242596.
PR 24-NOV-1998; 98JP-00333284.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Yokoyama S, Hirao I, Sakamoto K;
XX
DR WPI; 2000-224330/19.
XX
PT Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT target protein like Raf-1, useful in drug compositions to treat and
PT diagnose malignant tumors and inflammation.
XX
PS Disclosure; Page 52; 59pp; Japanese.
XX
CC The invention relates to novel nucleic acids which bind specifically to
CC the target protein of Ras, e.g. Raf-1. RNA sequences (AAZ99076-Z99077)
CC represent the sequence of aptamer inserts that bind the Ras target
CC protein which are used to generate PCR primers. Aptamers based on these
CC sequences are useful in the treatment and diagnosis of malignant tumours
CC and inflammation. The nucleic acids can be used to formulate medicinal
CC compositions that are useful in the treatment of malignant tumours and
CC inflammation and for disease diagnosis by binding specifically to Ras
CC target protein and regulating transmission of signal causing
CC proliferation or differentiation of cells
XX
SQ Sequence 45 BP; 11 A; 10 C; 7 G; 0 T; 17 U; 0 Other;

Query Match 45.0%; Score 45; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         27 AUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
              |||||||||||||||||||||||||||||||||||||||||||||
Db          1 AUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 45
```

RESULT 10
AAC75113/c
ID AAC75113 standard; cDNA; 748 BP.
XX
AC AAC75113;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF668 polynucleotide sequence SEQ ID NO:1335.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB40904.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 1146-1147; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 748 BP; 134 A; 218 C; 221 G; 174 T; 0 U; 1 Other;

Query Match 30.6%; Score 30.6; DB 3; Length 748;
Best Local Similarity 46.8%; Pred. No. 1.6;
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
Db 317 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 258

Qy 78 CCUUAGCGACAGCAAGC 94
Db 257 CCCTGGGCACGGGAAGC 241

RESULT 11
ABN24508/c
ID ABN24508 standard; cDNA; 748 BP.
XX
AC ABN24508;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX polynucleotide sequence SEQ ID NO:17493.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR P-PSDB; ABP08756.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 17493; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 748 BP; 134 A; 218 C; 221 G; 174 T; 0 U; 1 Other;

Query Match 30.6%; Score 30.6; DB 6; Length 748;
Best Local Similarity 46.8%; Pred. No. 1.6;
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

```
Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              :| ||   |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db        317 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 258

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db        257 CCCTGGGCACGGGAAGC 241
```

RESULT 12
AAS59583
ID AAS59583 standard; DNA; 15776 BP.
XX
AC AAS59583;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein encoding DNA #78.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Claim 1; SEQ ID NO 78; 1069pp; English.
XX
CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by P. acnes. The disorders include
CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of P. acnes in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for P. acnes proteins. These antibodies can be
CC used to downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU57347-AAU57508 and AAU67590-AAU67591. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 15776 BP; 3393 A; 4931 C; 4483 G; 2969 T; 0 U; 0 Other;

Query Match 30.6%; Score 30.6; DB 4; Length 15776;
Best Local Similarity 46.8%; Pred. No. 4;
Matches 36; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

```
Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              :| ||   |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db       8427 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 8486

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db       8487 CCCTGGGCACGGGAAGC 8503
```

RESULT 13
ACF64512
ID ACF64512 standard; DNA; 15776 BP.
XX
AC ACF64512;
XX
DT 17-OCT-2003 (first entry)
XX
DE Propionibacterium acnes DNA contig sequence #78.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; ds.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Claim 1; SEQ ID NO 78; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present

```
CC   sequence represents a P. acnes DNA contig which is specifically claimed
CC   in the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 15776 BP; 3393 A; 4931 C; 4483 G; 2969 T; 0 U; 0 Other;

  Query Match             30.6%;  Score 30.6;  DB 8;  Length 15776;
  Best Local Similarity   46.8%;  Pred. No. 4;
  Matches   36;  Conservative   12;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         18 UCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUG 77
              : | ||    |  :|||| ||  | : |::|| |||  : |: ||  ||:: :||   |:|
Db       8427 TTGCGGTCGACCTCGACGCCAGCGTTCTTCAGACCGTTGATGAAACGGTTGTAGGTCATG 8486

Qy         78 CCUUAGCGACAGCAAGC 94
              || : |  || | ||||
Db       8487 CCCTGGGCACGGGAAGC 8503


RESULT 14
ADA29321/c
ID   ADA29321 standard; DNA; 366 BP.
XX
AC   ADA29321;
XX
DT   20-NOV-2003  (first entry)
XX
DE   DNA encoding Acinetobacter baumannii protein #608.
XX
KW   ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW   vaccine; plant biocontrol agent.
XX
OS   Acinetobacter baumannii.
XX
PN   US6562958-B1.
XX
PD   13-MAY-2003.
XX
PF   04-JUN-1999;   99US-00328352.
XX
PR   09-JUN-1998;   98US-0088701P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton G,  Bush D;
XX
DR   WPI; 2003-576092/54.
DR   P-PSDB; ADA33447.
XX
PT   New Acinetobacter baumanii proteins and nucleic acids, useful as reagents
PT   for diagnosing a bacterial disease, as components of antibacterial
PT   vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT   plants.
XX
PS   Example; SEQ ID NO 608; 328pp; English.
XX
CC   The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC   The A. baumannii nucleic acids and polypeptides are useful as reagents
CC   for diagnosing a bacterial disease, as components of antibacterial
CC   vaccines, as targets for antibacterial drugs, to detect the presence of
CC   A. baumannii and other Acinetobacter species in a sample, in screening
CC   compounds for the ability to interfere with the A. baumannii life cycle
CC   or to inhibit A. baumannii infection, and as biocontrol agents for
CC   plants. The present sequence represents DNA encoding an A. baumannii
CC   protein.
XX
SQ   Sequence 366 BP; 131 A; 65 C; 63 G; 107 T; 0 U; 0 Other;

  Query Match             30.4%;  Score 30.4;  DB 9;  Length 366;
  Best Local Similarity   39.8%;  Pred. No. 1.5;
  Matches   35;  Conservative   17;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||| |:  | |||:  | |      |:|:|:||   | |:  :  : ||||| ||: |:|
Db        101 GGGTGTTTATGAATATAGCTTTCTATATGTCGCACCAGTAATTGTGCAAACAAGTAATAG 42

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACA 88
                :||:::  || :|:    :||   ||
Db         41 TCTGGTTTATGCGTATATTCTAGATTCA 14


RESULT 15
AAZ99049
ID   AAZ99049 standard; RNA; 98 BP.
XX
AC   AAZ99049;
XX
DT   21-JUN-2000  (first entry)
XX
DE   RNA aptamer #26 for binding Ras target protein.
XX
KW   Ras target protein; malignant tumour; signal transmission regulation;
KW   cell proliferation; cell differentiation; aptamer; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009684-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1999;   99WO-JP004399.
XX
PR   14-AUG-1998;   98JP-00242596.
PR   24-NOV-1998;   98JP-00333284.
XX
PA   (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI   Yokoyama S,  Hirao I,  Sakamoto K;
XX
DR   WPI; 2000-224330/19.
XX
PT   Nucleic acid e.g. RNA aptamer capable of binding specifically to Ras
PT   target protein like Raf-1, useful in drug compositions to treat and
PT   diagnose malignant tumors and inflammation.
XX
PS   Claim 6; Page 41; 59pp; Japanese.
XX
CC   The invention relates to novel nucleic acids which bind specifically to
CC   the target protein of Ras, e.g. Raf-1. RNA aptamer (AAZ99024-Z99051)
CC   based on these sequences are useful in the treatment and diagnosis of
CC   malignant tumours and inflammation. The nucleic acids can be used to
CC   formulate medicinal compositions that are useful in the treatment of
CC   malignant tumours and inflammation and for disease diagnosis by binding
CC   specifically to Ras target protein and regulating transmission of signal
CC   causing proliferation or differentiation of cells
XX
SQ   Sequence 98 BP; 24 A; 24 C; 26 G; 0 T; 24 U; 0 Other;

  Query Match             30.0%;  Score 30;  DB 3;  Length 98;
  Best Local Similarity   63.0%;  Pred. No. 1.3;
  Matches   63;  Conservative    0;  Mismatches   35;  Indels    2;  Gaps    1;

Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              |||||||||||||||||||||||||        |   |   |       ||  | |
Db          1 GGGAGUGGAGGAAUUCAUCGAGGCAUGACCUCCCGUGGCAGUAGGGGUAAAAAUUAUCUU 60

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUGC 100
                |    ||  |  ||||||||||||||||||||||||||
Db         61 CCUACACUU--CUCAUGCCUUAGCGACAGCAAGCUUCUGC 98


Search completed: November 30, 2004, 11:19:22
Job time : 237.222 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:16 ; Search time 51.1111 Seconds
(without alignments)
1390.674 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaauucaucg.....uagcgacagcaagcuucugc 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	30.4	30.4	366	4	US-09-328-352-608	Sequence 608, App
2	29.8	29.8	4285	3	US-09-410-464-1	Sequence 1, Appli
3	27.4	27.4	405	4	US-09-270-767-6441	Sequence 6441, Ap
4	27.4	27.4	405	4	US-09-270-767-21723	Sequence 21723, A
c 5	27.4	27.4	3791	3	US-08-675-773B-3	Sequence 3, Appli
c 6	27	27.0	208	4	US-09-513-999C-21500	Sequence 21500, A
c 7	26.8	26.8	415	4	US-09-513-999C-26061	Sequence 26061, A
8	25.8	25.8	798	4	US-08-956-171E-162	Sequence 162, App
9	25.8	25.8	798	4	US-08-781-986A-162	Sequence 162, App
10	25.6	25.6	372	4	US-08-956-171E-798	Sequence 798, App
11	25.6	25.6	372	4	US-08-781-986A-798	Sequence 798, App
12	25.2	25.2	355	4	US-09-513-999C-8830	Sequence 8830, Ap
c 13	25.2	25.2	786431	4	US-09-751-389-3	Sequence 3, Appli
14	25	25.0	246	4	US-09-513-999C-35958	Sequence 35958, A
c 15	25	25.0	960	4	US-09-328-352-1916	Sequence 1916, Ap
16	25	25.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
17	25	25.0	1830121	4	US-09-643-990A-1	Sequence 1, Appli
18	25	25.0	1830121	4	US-10-329-960-1	Sequence 1, Appli
19	24.8	24.8	2775	4	US-09-248-796A-4695	Sequence 4695, Ap
20	24.8	24.8	3786	3	US-08-975-762-42	Sequence 42, Appl
21	24.8	24.8	3786	3	US-09-295-028-42	Sequence 42, Appl
22	24.8	24.8	3786	3	US-09-106-582-42	Sequence 42, Appl
23	24.8	24.8	3786	4	US-09-159-469-42	Sequence 42, Appl
24	24.8	24.8	3786	4	US-09-693-542-42	Sequence 42, Appl
25	24.8	24.8	7399	2	US-08-418-848A-9	Sequence 9, Appli
26	24.8	24.8	9709	2	US-08-188-583-5	Sequence 5, Appli
27	24.8	24.8	9709	3	US-08-388-353-1	Sequence 1, Appli
28	24.8	24.8	9709	3	US-08-488-551B-1	Sequence 1, Appli
29	24.8	24.8	9709	4	US-09-309-572-15	Sequence 15, Appl
30	24.8	24.8	9709	4	US-09-718-096-15	Sequence 15, Appl
31	24.8	24.8	12479	4	US-09-318-138-13	Sequence 13, Appl
32	24.8	24.8	12494	3	US-08-935-312-13	Sequence 13, Appl
33	24.8	24.8	12494	3	US-08-848-760B-33	Sequence 33, Appl
34	24.8	24.8	15581	3	US-08-646-538-35	Sequence 35, Appl
35	24.8	24.8	15581	3	US-09-503-222-35	Sequence 35, Appl
c 36	24.6	24.6	544	4	US-09-621-976-1765	Sequence 1765, Ap
37	24.6	24.6	1212	4	US-09-248-796A-136	Sequence 136, App
38	24.6	24.6	1827	4	US-09-270-767-29735	Sequence 29735, A
39	24.6	24.6	2612	3	US-09-105-390-7	Sequence 7, Appli
40	24.6	24.6	4285	4	US-09-270-767-13714	Sequence 13714, A
41	24.6	24.6	11802	4	US-08-956-171E-70	Sequence 70, Appl
42	24.6	24.6	11802	4	US-08-781-986A-70	Sequence 70, Appl
c 43	24.4	24.4	915	4	US-09-919-497-19	Sequence 19, Appl
44	24.4	24.4	13542	4	US-08-956-171E-154	Sequence 154, App
45	24.4	24.4	13542	4	US-08-781-986A-154	Sequence 154, App

ALIGNMENTS

RESULT 1
US-09-328-352-608/c
; Sequence 608, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 608
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-608

Query Match 30.4%; Score 30.4; DB 4; Length 366;
Best Local Similarity 39.8%; Pred. No. 0.088;
Matches 35; Conservative 17; Mismatches 36; Indels 0; Gaps 0;

```
Qy          1 GGGAGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAA 60
              ||| |:  | |||:  | |     |:|:|:||   | |:  :  : ||||| ||: |:|
Db        101 GGGTGTTTATGAATATAGCTTTCTATATGTCGCACCAGTAATTGTGCAAACAAGTAATAG 42

Qy         61 AUUGGUUUUAGCAUAUGCCUUAGCGACA 88
                :||:::  || :|:    :||   ||
Db         41 TCTGGTTTATGCGTATATTCTAGATTCA 14
```

RESULT 2
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4285

```
;   TYPE: DNA
;   ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-1

  Query Match             29.8%;  Score 29.8;  DB 3;  Length 4285;
  Best Local Similarity   33.3%;  Pred. No. 0.33;
  Matches   27;  Conservative   22;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              ||   | :| :  |  | : : :||  :   : : ||::||| |||  :|:  |::| ::
Db       1193 GATTTAATCCTGTAACCTTCTTTCGTTTTTTTATGCCTTCAATCCATCTATTTATTGTTT 1252

Qy         68 UUAGCAUAUGCCUUAGCGACA 88
              ::|  |: :  |::||  |||
Db       1253 TTATGATTTTTCTTAGATACA 1273


RESULT 3
US-09-270-767-6441
; Sequence 6441, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6441
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-6441

  Query Match             27.4%;  Score 27.4;  DB 4;  Length 405;
  Best Local Similarity   35.5%;  Pred. No. 1.2;
  Matches   33;  Conservative   19;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          7 GGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGU 66
              || | ||: ||:   |    |: : |  :|| :|:    :|||||| | :|  ||:   :
Db        227 GGTGTAATCCATTTGGTTCAATTTTGCGTCCTTCTCAACTCAAACCTGGTAATAATCCAT 286

Qy         67 UUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUG 99
              :::   |  :  |: | | |    ||| ::|:|
Db        287 TTTTAAAAGTCGCTGAACCAACCGAAGATTCTG 319


RESULT 4
US-09-270-767-21723
; Sequence 21723, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21723
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-21723

  Query Match             27.4%;  Score 27.4;  DB 4;  Length 405;
  Best Local Similarity   35.5%;  Pred. No. 1.2;
  Matches   33;  Conservative   19;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          7 GGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGU 66
              || | ||: ||:   |    |: : |  :|| :|:    :|||||| | :|  ||:   :
Db        227 GGTGTAATCCATTTGGTTCAATTTTGCGTCCTTCTCAACTCAAACCTGGTAATAATCCAT 286

Qy         67 UUUAGCAUAUGCCUUAGCGACAGCAAGCUUCUG 99
              :::   |  :  |: | | |    ||| ::|:|
Db        287 TTTTAAAAGTCGCTGAACCAACCGAAGATTCTG 319


RESULT 5
US-08-675-773B-3/c
; Sequence 3, Application US/08675773B
; Patent No. 6166288
;  GENERAL INFORMATION:
;    APPLICANT:  DIAMOND, LISA E
;    APPLICANT:  LOGAN, JOHN S
;    APPLICANT:  BYRNE, GUERARD W
;    APPLICANT:  SHARMA, AJAY
;    TITLE OF INVENTION:  METHOD OF PRODUCING TRANSGENIC ANIMALS
;    TITLE OF INVENTION:  FOR XENOTRANSPLANTATION. . .
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  419 SEVENTH STREET, N.W., SUITE 300
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/675,773B
;      FILING DATE:  03-JUL-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, IVER P.
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  DIAMOND=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-675-773B-3

  Query Match             27.4%;  Score 27.4;  DB 3;  Length 3791;
  Best Local Similarity   41.6%;  Pred. No. 2.5;
  Matches   32;  Conservative   14;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         19 CGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGC 78
              | |||  :  |  ||| | |:|: ||:  |  ||  ::|: |  :| : : ||||: : |
Db        624 CCAGGATTCAGCAGACCCTGTCTCCCTGAACCCCTCTTATTAGATGTTGTGAGCATTTTC 565

Qy         79 CUUAGCGACAGCAAGCU 95
              | : | |||   | ||:
Db        564 CATGGGGACTTGATGCT 548


RESULT 6
US-09-513-999C-21500/c
; Sequence 21500, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
```

```
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 21500
;   LENGTH: 208
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 13
;   OTHER INFORMATION: v=a or c or g
US-09-513-999C-21500

  Query Match             27.0%;  Score 27;  DB 4;  Length 208;
  Best Local Similarity   39.8%;  Pred. No. 1.3;
  Matches   33;  Conservative   15;  Mismatches   35;  Indels    0;  Gaps    0;

Qy         13 AUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGC 72
              |:: |    |  ||  :|:| | :| | |::     || |||| :  | |:: | :: | |
Db        203 ATTAACACAACCAACTGTCAAGTCTGCCTTTAAAAAACCCAGCTTGACATTTGCTTCATC 144

Qy         73 AUAUGCCUUAGCGACAGCAAGCU 95
              |:|: |      |  |||| |||:
Db        143 ATATTCAGACCCCCCAGCTAGCT 121


RESULT 7
US-09-513-999C-26061/c
; Sequence 26061, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 26061
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 340
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 344
;   OTHER INFORMATION: k=g or t
US-09-513-999C-26061

  Query Match             26.8%;  Score 26.8;  DB 4;  Length 415;
  Best Local Similarity   48.4%;  Pred. No. 2;
  Matches   30;  Conservative   10;  Mismatches   22;  Indels    0;  Gaps    0;

Qy         28 UGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGAC 87
              :|: ||| ||  |:  | : | |||||: |  | : ||:  :  ||:|: ||  || |||
Db        269 TGTAGACACCACCTGGCATTACACCAGTGAACACTAGGTGGTGCCATATCCCAAAGAGAC 210

Qy         88 AG 89
              ||
Db        209 AG 208


RESULT 8
US-08-956-171E-162
; Sequence 162, Application US/08956171E
; Patent No. 6593114
;  GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 162:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 798 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 162:
US-08-956-171E-162

  Query Match             25.8%;  Score 25.8;  DB 4;  Length 798;
  Best Local Similarity   31.2%;  Pred. No. 5.7;
  Matches   24;  Conservative   21;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         22 GGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUU 81
              ||||: : :|   : | :|::  ::     |    : :||| :||:::|| |:    | :
Db        196 GGCATCTTTCACTTTCTTCTTATTTTCGTCAGTATTTAAAGTGGTTTTAGGATTCCTCAT 255

Qy         82 AGCGACAGCAAGCUUCU 98
              ||  | || |  ::|:
Db        256 CGCTTCTGCGATTTTCT 272


RESULT 9
US-08-781-986A-162
; Sequence 162, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Human Genome Sciences, Inc.
;       STREET:  9410 Key West Avenue
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;       COMPUTER:  HP Vectra 486/33
;       OPERATING SYSTEM:  MSDOS version 6.2
;       SOFTWARE:  ASCII Text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/781,986A
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Benson, Bob
;       REGISTRATION NUMBER:  30,446
;       REFERENCE/DOCKET NUMBER:  PB248PP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 309-8504
;       TELEFAX:  (301) 309-8512
;   INFORMATION FOR SEQ ID NO:  162:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  798 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-781-986A-162

  Query Match             25.8%;  Score 25.8;  DB 4;  Length 798;
  Best Local Similarity   31.2%;  Pred. No. 5.7;
  Matches   24;  Conservative   21;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         22 GGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUU 81
              ||||: : :|   : | :|::  ::    |    : :||| :||:::::|| |:    | :
Db        196 GGCATCTTTCACTTTCTTCTTATTTTCGTCAGTATTTAAAGTGGTTTTAGGATTCCTCAT 255

Qy         82 AGCGACAGCAAGCUUCU 98
                ||  | || |  ::|:
Db        256 CGCTTCTGCGATTTTCT 272


RESULT 10
US-08-956-171E-798
; Sequence 798, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 798:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 372 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 798:
US-08-956-171E-798

  Query Match             25.6%;  Score 25.6;  DB 4;  Length 372;
  Best Local Similarity   36.9%;  Pred. No. 5.3;
  Matches   24;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         27 AUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGA 86
              |: :| || || : ::| ::|| ||| :: :||   | : :  |||: :|  ::  |||
Db        223 ATTTCAACACCTTTTTCTTTCATACCTTTTTTAACAGGTTGTGTGCATTTGNTTTTTCGA 282

Qy         87 CAGCA 91
                ||
Db        283 GGCCA 287


RESULT 11
US-08-781-986A-798
; Sequence 798, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  798:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  372 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-798

  Query Match             25.6%;  Score 25.6;  DB 4;  Length 372;
  Best Local Similarity   36.9%;  Pred. No. 5.3;
  Matches   24;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         27 AUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGA 86
              |: :| || || : ::| ::|| ||| :: :||   | : :  |||: :|  ::  |||
Db        223 ATTTCAACACCTTTTTCTTTCATACCTTTTTTAACAGGTTGTGTGCATTTGNTTTTTCGA 282

Qy         87 CAGCA 91
                 ||
Db        283 GGCCA 287


RESULT 12
US-09-513-999C-8830
; Sequence 8830, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 8830
;   LENGTH: 355
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 299
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 309
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 319
;   OTHER INFORMATION: r=a or g
US-09-513-999C-8830

  Query Match             25.2%;  Score 25.2;  DB 4;  Length 355;
  Best Local Similarity   37.1%;  Pred. No. 7.3;
  Matches   23;  Conservative   16;  Mismatches   23;  Indels    0;  Gaps    0;

Qy         27 AUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGA 86
              |:  | | :||   ::   : ||||| |:: :   ::||::::  ||:  || :  ||||
Db        166 ATTACAATTCCTGTTTTTCTAAAACCGGTTTTTCTTTGGTTTTGTCATTAGCATGTGCGA 225

Qy         87 CA 88
              ||
Db        226 CA 227


RESULT 13
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
;  APPLICANT: GUEGLER, Karl et al
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001067
;  CURRENT APPLICATION NUMBER: US/09/751,389
;  CURRENT FILING DATE:  2001-01-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 786431
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(786431)
;   OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

  Query Match             25.2%;  Score 25.2;  DB 4;  Length 786431;
  Best Local Similarity   42.6%;  Pred. No. 91;
  Matches   23;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy         44 CUUCAAACCAGUUAUAAAUUGGUUUUAGCAUAUGCCUUAGCGACAGCAAGCUUC 97
              |: || | | |: |:| |::| ::: |||   :| |::|| ||||    | ::|
Db     507097 CTCCAGAGCTGTGATATATTGTTTTAAGCCATTGACTTAGTGACAATTTGTTTC 507044


RESULT 14
US-09-513-999C-35958
; Sequence 35958, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 35958
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 137
;   OTHER INFORMATION: m=a or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 143
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 153
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 170
;   OTHER INFORMATION: b=c or g or t
US-09-513-999C-35958

  Query Match             25.0%;  Score 25;  DB 4;  Length 246;
  Best Local Similarity   43.9%;  Pred. No. 7.7;
  Matches   18;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         30 UCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUA 70
              : | |:  | | :||::||  ||||: |:|| ::||::::|
Db        182 TAGTCTGGGGCATCCTTCATGCCAGTGATAATTTGGTTTTA 222
```

```
RESULT 15
US-09-328-352-1916/c
; Sequence 1916, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1916
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-1916

  Query Match             25.0%;  Score 25;  DB 4;  Length 960;
  Best Local Similarity   36.9%;  Pred. No. 12;
  Matches   24;  Conservative   16;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         12 AAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
              ||::  :| ||  |:| | ||||:   :| : |::||    | : |:|||:    :::
Db        534 AATTGCTCTAGCAATACGGCGACTATATCGTTCTTCACCATATTGATAAATGATATTTGC 475

Qy         72 CAUAU 76
              || |:
Db        474 CAAAT 470


Search completed: November 30, 2004, 11:53:59
Job time : 58.1111 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 30, 2004, 11:12:16 ; Search time 231.667 Seconds
(without alignments)
2339.337 Million cell updates/sec

Title: US-09-529-397C-25
Perfect score: 100
Sequence: 1 gggaguggaggaauucaucg.....uagcgacagcaagcuucugc 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3640242 seqs, 2709731945 residues

Total number of hits satisfying chosen parameters: 7280484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	29.8	29.8	4285	14	US-10-104-580-1	Sequence 1, Appli
c	2	28.8	28.8	675	16	US-10-282-122A-40631	Sequence 40631, A
c	3	28.6	28.6	2000	9	US-09-938-842A-2908	Sequence 2908, Ap
c	4	28.6	28.6	2000	11	US-09-938-842A-2908	Sequence 2908, Ap
c	5	28.2	28.2	586	16	US-10-425-114-30801	Sequence 30801, A
c	6	28.2	28.2	1612	17	US-10-767-701-13576	Sequence 13576, A
c	7	28.2	28.2	1851	18	US-10-425-115-121623	Sequence 121623,
c	8	28	28.0	1691139	14	US-10-067-514-1	Sequence 1, Appli
c	9	28	28.0	1691139	16	US-10-419-723-1	Sequence 1, Appli
	10	27.8	27.8	1093	15	US-10-369-493-25760	Sequence 25760, A
c	11	27.8	27.8	99291	17	US-10-322-281-744	Sequence 744, App
	12	27	27.0	522	15	US-10-097-111-47	Sequence 47, Appl
	13	27	27.0	56506	15	US-10-097-111-10	Sequence 10, Appl
	14	26.8	26.8	1116	15	US-10-271-181B-6	Sequence 6, Appli
	15	26.6	26.6	1884	15	US-10-369-493-35948	Sequence 35948, A
c	16	26.6	26.6	48995	16	US-10-052-482-85	Sequence 85, Appl
c	17	26.4	26.4	621	13	US-10-027-632-199104	Sequence 199104,
c	18	26.4	26.4	621	15	US-10-027-632-199104	Sequence 199104,
c	19	26.2	26.2	489	10	US-09-814-353-15351	Sequence 15351, A
c	20	26.2	26.2	502	13	US-10-027-632-324279	Sequence 324279,
c	21	26.2	26.2	502	13	US-10-027-632-324280	Sequence 324280,
c	22	26.2	26.2	502	15	US-10-027-632-324279	Sequence 324279,
c	23	26.2	26.2	502	15	US-10-027-632-324280	Sequence 324280,
	24	26.2	26.2	181684	13	US-10-087-192-790	Sequence 790, App
c	25	26	26.0	368	16	US-10-424-599-130987	Sequence 130987,
	26	26	26.0	725	13	US-10-027-632-150280	Sequence 150280,
	27	26	26.0	725	15	US-10-027-632-150280	Sequence 150280,
	28	26	26.0	740	13	US-10-027-632-14431	Sequence 14431, A
	29	26	26.0	740	13	US-10-027-632-14432	Sequence 14432, A
	30	26	26.0	740	15	US-10-027-632-14431	Sequence 14431, A
	31	26	26.0	740	15	US-10-027-632-14432	Sequence 14432, A
	32	26	26.0	2515	13	US-10-027-632-101724	Sequence 101724,
	33	26	26.0	2515	15	US-10-027-632-101724	Sequence 101724,
	34	26	26.0	7286	14	US-10-198-846-13160	Sequence 13160, A
	35	26	26.0	8998	15	US-10-311-455-1680	Sequence 1680, Ap
	36	25.8	25.8	798	8	US-08-781-986A-162	Sequence 162, App
	37	25.8	25.8	798	16	US-10-329-624-162	Sequence 162, App
c	38	25.8	25.8	1165	18	US-10-425-115-310	Sequence 310, App
c	39	25.8	25.8	1550	17	US-10-437-963-73812	Sequence 73812, A
c	40	25.8	25.8	2577	17	US-10-437-963-73810	Sequence 73810, A
c	41	25.8	25.8	2688	17	US-10-437-963-73813	Sequence 73813, A
	42	25.8	25.8	25320	13	US-10-087-192-253	Sequence 253, App
	43	25.8	25.8	43980	16	US-10-398-221-5	Sequence 5, Appli
	44	25.8	25.8	94330	13	US-10-087-192-1222	Sequence 1222, Ap
c	45	25.8	25.8	226215	13	US-10-087-192-1948	Sequence 1948, Ap

ALIGNMENTS

RESULT 1
US-10-104-580-1
; Sequence 1, Application US/10104580
; Publication No. US20030033628A1
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 62486
; CURRENT APPLICATION NUMBER: US/10/104,580
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/410,464
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/287,700
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: 60/080,851
; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4285
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
US-10-104-580-1

Query Match 29.8%; Score 29.8; DB 14; Length 4285;
Best Local Similarity 33.3%; Pred. No. 9.7;
Matches 27; Conservative 22; Mismatches 32; Indels 0; Gaps 0;

```
Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              ||   | :|  :   |   | : : :||   :    : :  ||::||| |||   :|:   |::| ::
Db       1193 GATTTAATCCTGTAACCTTCTTTCGTTTTTTATGCCTTCAATCCATCTATTTATTGTTT 1252

Qy         68 UUAGCAUAUGCCUUAGCGACA 88
                ::|  |: :  |::||  |||
```

```
Db       1253 TTATGATTTTTCTTAGATACA 1273


RESULT 2
US-10-282-122A-40631/c
; Sequence 40631, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40631
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-10-282-122A-40631

  Query Match             28.8%;  Score 28.8;  DB 16;  Length 675;
  Best Local Similarity   42.2%;  Pred. No. 11;
  Matches   27;  Conservative   15;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          4 AGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUU 63
              ||:|||| ||::   |||| ||:|:  |||      | : |:: ||    |::|:| |::
Db        563 AGTGGAGCAATTTTGCGAGTCATATCACGATGATCACGTTCTTTAATATCGTTATATATT 504

Qy         64 GGUU 67
              | ::
Db        503 GATT 500


RESULT 3
US-09-938-842A-2908/c
; Sequence 2908, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2908
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2908

  Query Match             28.6%;  Score 28.6;  DB 9;  Length 2000;
  Best Local Similarity   35.8%;  Pred. No. 19;
  Matches   24;  Conservative   19;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 AGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUU 63
              ||:  | |||::|| | |   |: : :|| |    : ::|  : ||| |  ::|:| |::
Db        759 AGTACAAGAATTCAGCTACATATGTATCGCCGTAATTTTCTCTGAAAACTATTATATATT 700

Qy         64 GGUUUUA 70
              | ::::|
Db        699 GTTTTTA 693


RESULT 4
US-09-938-842A-2908/c
; Sequence 2908, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2908
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2908

  Query Match             28.6%;  Score 28.6;  DB 11;  Length 2000;
  Best Local Similarity   35.8%;  Pred. No. 19;
  Matches   24;  Conservative   19;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 AGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUU 63
              ||:  | |||::|| | |   |: : :|| |    : ::|  : ||| |  ::|:| |::
Db        759 AGTACAAGAATTCAGCTACATATGTATCGCCGTAATTTTCTCTGAAAACTATTATATATT 700

Qy         64 GGUUUUA 70
              | ::::|
Db        699 GTTTTTA 693
```

RESULT 5
US-10-425-114-30801/c
; Sequence 30801, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30801
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73080F09_FLI
US-10-425-114-30801

Query Match 28.2%; Score 28.2; DB 16; Length 586;
Best Local Similarity 37.1%; Pred. No. 17;
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

```
Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              | || |::  :       :|:   || :|  :| : |: ||   |  : |:|| : ||::
Db         92 GGGGCATTTTTATGTTGTTATTAGGATTCTCTCCTGCTCCACCGCCATCATAACTCGGTT 33

Qy         68 UUAGCAUAUGCCUUAGCGACAGCAAGCUU 96
                ||||   || ::|| || || ||| ::
Db         32 CCAGCAAGCGCATTAGGGAAAGGAAGATT 4
```

RESULT 6
US-10-767-701-13576/c
; Sequence 13576, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 13576
; LENGTH: 1612
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS1303_1
US-10-767-701-13576

Query Match 28.2%; Score 28.2; DB 17; Length 1612;
Best Local Similarity 38.2%; Pred. No. 24;
Matches 34; Conservative 17; Mismatches 38; Indels 0; Gaps 0;

```
Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              | || |::  :       :|:   || :|  :| : |: ||   | |: | || : ||::
Db       1152 GCGGCATTTTTGTGTTGTTATTAGGATTCTCTCCTGCTCCACTGCCGTCACAACTCGGTT 1093

Qy         68 UUAGCAUAUGCCUUAGCGACAGCAAGCUU 96
                ||||   || ::|| || || ||| ::
Db       1092 CCAGCAAGCGCATTAGGGAAAGGAAGATT 1064
```

RESULT 7
US-10-425-115-121623/c
; Sequence 121623, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 121623
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_42398C.1
US-10-425-115-121623

Query Match 28.2%; Score 28.2; DB 18; Length 1851;
Best Local Similarity 37.1%; Pred. No. 26;
Matches 33; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

```
Qy          8 GAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUU 67
              | || |::  :       :|:   || :|  :| : |: ||   |  : |:|| : ||::
Db       1333 GGGGCATTTTTATGTTGTTATTAGGATTCTCTCCTGCTCCACCGCCATCATAACTCGGTT 1274

Qy         68 UUAGCAUAUGCCUUAGCGACAGCAAGCUU 96
                ||||   || ::|| || || ||| ::
Db       1273 CCAGCAAGCGCATTAGGGAAAGGAAGATT 1245
```

RESULT 8
US-10-067-514-1/c
; Sequence 1, Application US/10067514
; Publication No. US20030054531A1
; GENERAL INFORMATION:
; APPLICANT: Gretarsdottir, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reynisdottir, Sigridur Th.
; TITLE OF INVENTION: HUMAN STROKE GENE
; FILE REFERENCE: 2345.2010-003
; CURRENT APPLICATION NUMBER: US/10/067,514
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 09/811/352
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
; TYPE: DNA
; ORGANISM: Human
US-10-067-514-1

Query Match 28.0%; Score 28; DB 14; Length 1691139;
Best Local Similarity 41.7%; Pred. No. 3.6e+02;
Matches 25; Conservative 15; Mismatches 20; Indels 0; Gaps 0;

```
Qy         12 AAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
              | :: |:| | || : :|  |    | : : ||::|||| |||: |:|||   |::::||
Db     127542 AGTTGATCCATGCTTTTGGAGGGAGCTTATCCCTTCAAAGCAGTAATAAAAGTGTTTTAG 127483
```

RESULT 9
US-10-419-723-1/c
; Sequence 1, Application US/10419723
; Publication No. US20040014099A1
; GENERAL INFORMATION:
; APPLICANT: Gretarsdottir, Solveig

```
;  APPLICANT:  Jonsdottir, Sif
;  APPLICANT:  Reynisdottir, Sigridur Th.
;  APPLICANT:  Thorleifsson, Gudmar
;  APPLICANT:  Gulcher, Jeffrey
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
;  TITLE OF INVENTION:  METHODS OF TREATMENT
;  FILE REFERENCE: 2345.2010-005
;  CURRENT APPLICATION NUMBER: US/10/419,723
;  CURRENT FILING DATE:  2003-04-18
;  PRIOR APPLICATION NUMBER: 10/255,120
;  PRIOR FILING DATE: 2002-09-25
;  PRIOR APPLICATION NUMBER: 10/067,514
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 09/811,352
;  PRIOR FILING DATE: 2001-03-19
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1691139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-419-723-1

  Query Match             28.0%;  Score 28;  DB 16;  Length 1691139;
  Best Local Similarity   41.7%;  Pred. No. 3.6e+02;
  Matches   25;  Conservative   15;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         12 AAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
              | :: |:| | || : :|  |    | : : ||::|||| |||: |:|||   |::::||
Db     127542 AGTTGATCCATGCTTTTGGAGGGAGCTTATCCCTTCAAAGCAGTAATAAAAGTGTTTTAG 127483


RESULT 10
US-10-369-493-25760
; Sequence 25760, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25760
;   LENGTH: 1093
;   TYPE: DNA
;   ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25760

  Query Match             27.8%;  Score 27.8;  DB 15;  Length 1093;
  Best Local Similarity   38.0%;  Pred. No. 29;
  Matches   27;  Conservative   17;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          4 AGUGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUU 63
              ||:|||    ::||: ||     | |  |  :|||: : ||::  ||| ||:: :|||:
Db        213 AGTGGACTCTTTCATAGACAATGAAGCAGGATCCGTTTACCTTATAACGAGTTTTAAATC 272

Qy         64 GGUUUUAGCAU 74
              | :: :   |:
Db        273 GTTTGTTCTAT 283


RESULT 11
US-10-322-281-744/c
; Sequence 744, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001000
;  CURRENT APPLICATION NUMBER: US/10/322,281
;  CURRENT FILING DATE:  2002-12-17
;  NUMBER OF SEQ ID NOS: 866
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 744
;   LENGTH: 99291
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-322-281-744

  Query Match             27.8%;  Score 27.8;  DB 17;  Length 99291;
  Best Local Similarity   31.0%;  Pred. No. 1.6e+02;
  Matches   22;  Conservative   22;  Mismatches   27;  Indels    0;  Gaps    0;

Qy         12 AAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAG 71
              ||:  |: ||| | : : :|  : | : ::  ::   | ||  :|:||| :||:::|
Db      23527 AATAAATTGAGACCTGTCTCAGATACTTTTGGTTTGCAACAAATATAAAATGGTTTTAC 23468

Qy         72 CAUAUGCCUUA 82
              |: :|  ::|
Db      23467 AATTTGGTTTA 23457


RESULT 12
US-10-097-111-47
; Sequence 47, Application US/10097111
; Publication No. US20030138771A1
; GENERAL INFORMATION:
;  APPLICANT: PELLETIER, JERRY
;  APPLICANT:  GROS, PHILLIPPE
;  APPLICANT:  DUBOW, MICHAEL
;  TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT
;  TITLE OF INVENTION:  ENCODE ANTI-MICROBIAL POLYPEPTIDES
;  FILE REFERENCE: 073406-0603
;  CURRENT APPLICATION NUMBER: US/10/097,111
;  CURRENT FILING DATE:  2002-07-24
;  PRIOR APPLICATION NUMBER: 09/676,412
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/157,218
;  PRIOR FILING DATE: 1999-09-30
;  NUMBER OF SEQ ID NOS: 552
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;   LENGTH: 522
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-10-097-111-47

  Query Match             27.0%;  Score 27;  DB 15;  Length 522;
  Best Local Similarity   40.0%;  Pred. No. 42;
  Matches   30;  Conservative   15;  Mismatches   30;  Indels    0;  Gaps    0;

Qy         13 AUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGC 72
              |::||: ||   |  :| | || | |:  : |::|||   | : |   | : | :::|||
Db        186 ATTCATTGACAGACTTGACCACGCTGTTCTTCTTCAAGGGAATGAACCAATCGCTTTAGC 245

Qy         73 AUAUGCCUUAGCGAC 87
              | |:||  :  |  ||
Db        246 AAATGCAGTTGACAC 260


RESULT 13
US-10-097-111-10
; Sequence 10, Application US/10097111
; Publication No. US20030138771A1
; GENERAL INFORMATION:
;  APPLICANT: PELLETIER, JERRY
;  APPLICANT:  GROS, PHILLIPPE
```

```
;  APPLICANT:  DUBOW, MICHAEL
;  TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT
;  TITLE OF INVENTION:  ENCODE ANTI-MICROBIAL POLYPEPTIDES
;  FILE REFERENCE: 073406-0603
;  CURRENT APPLICATION NUMBER: US/10/097,111
;  CURRENT FILING DATE:  2002-07-24
;  PRIOR APPLICATION NUMBER: 09/676,412
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/157,218
;  PRIOR FILING DATE: 1999-09-30
;  NUMBER OF SEQ ID NOS: 552
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 56506
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-10-097-111-10

  Query Match             27.0%;  Score 27;  DB 15;  Length 56506;
  Best Local Similarity   40.0%;  Pred. No. 2.4e+02;
  Matches   30;  Conservative   15;  Mismatches   30;  Indels    0;  Gaps    0;

Qy         13 AUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGC 72
              |::||: ||   |  :| | || | |:  : |::|||   | : |   | : | :::|||
Db       1535 ATTCATTGACAGACTTGACCACGCTGTTCTTCTTCAAGGGAATGAACCAATCGCTTTAGC 1594

Qy         73 AUAUGCCUUAGCGAC 87
              | |:||  : |  ||
Db       1595 AAATGCAGTTGACAC 1609


RESULT 14
US-10-271-181B-6
; Sequence 6, Application US/10271181B
; Publication No. US20030158672A1
; GENERAL INFORMATION:
;  APPLICANT: Kalyanaraman Ramnarayan
;  APPLICANT:  Edward T. Maggio
;  APPLICANT:  P. Patrick Hess
;  TITLE OF INVENTION: Use of Computationally Derived Protein
;  TITLE OF INVENTION:  Structures of Genetic Polymorphisms in Pharmacogenomics for
;  TITLE OF INVENTION:  Drug Design and Clinical Applications
;  FILE REFERENCE: 24737-1906D
;  CURRENT APPLICATION NUMBER: US/10/271,181B
;  CURRENT FILING DATE:  2002-10-10
;  PRIOR APPLICATION NUMBER: 09/438,566
;  PRIOR FILING DATE: 1999-11-10
;  PRIOR APPLICATION NUMBER: 09/704,362
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/709,905
;  PRIOR FILING DATE: 2000-11-10
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 1116
;   TYPE: DNA
;   ORGANISM: Human Immunodeficiency Virus (HIV)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(297)
;   OTHER INFORMATION: HIV Protease
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (298)...(1116)
;   OTHER INFORMATION: Portion of HIV Reverse Transcriptase
US-10-271-181B-6

  Query Match             26.8%;  Score 26.8;  DB 15;  Length 1116;
  Best Local Similarity   38.4%;  Pred. No. 65;
  Matches   33;  Conservative   16;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          6 UGGAGGAAUUCAUCGAGGCAUAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGG 65
              : |||   ::||   |   | |:   |||   |:| :|  :|||  ||   |: | ::|
Db        800 TAGAGCCTTTCAGAAAACAAAATCCAGACATGGTCATCTATCAATACATGGATGATTTGT 859

Qy         66 UUUUAGCAUAUGCCUUAGCGACAGCA 91
                : :|| |: :| |::||  | || |
Db        860 ATGTAGGATCTGACTTAGAAATAGGA 885


RESULT 15
US-10-369-493-35948
; Sequence 35948, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35948
;   LENGTH: 1884
;   TYPE: DNA
;   ORGANISM: Mesorhizobium loti
US-10-369-493-35948

  Query Match             26.6%;  Score 26.6;  DB 15;  Length 1884;
  Best Local Similarity   49.0%;  Pred. No. 94;
  Matches   24;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         26 UAUGUCGACUCCGUCUUCCUUCAAACCAGUUAUAAAUUGGUUUUAGCAU 74
              :|:| |||  |||:|  || :|||||| | :|: ||::|  : : || :
Db        364 TATGGCGAGGCCGTCGGCCATCAAACCGGCTATGAATTGCGTCTCGCCT 412


Search completed: November 30, 2004, 12:01:02
Job time : 237.667 secs
```